Errata

p.34	line 7,	object by mapping
p.43	line 15,	than E
p.48	line 17,	suggested by Gabor
p.57	line 8,	reduced contrast
p.57	line 9,	therefore T(u)
p.63	line 10,	be remembered
p.69	line 8	fringes is highest
p.95	line 20	a typical point
p.101	line 2	microscope objective D
p.111	line 27	same conditions
p.216	line 4	Phot. Sci & Eng 10, 133 (1966)

UNIVERSITY OF LONDON

IMPERIAL COLLEGE OF SCIENCE AND TECHNOLOGY

Applied Optics Section

Department of Physics

A STUDY OF SOME OPTICAL PROPERTIES OF PHOTOGRAPHIC EMULSIONS, WITH PARTICULAR REFERENCE TO RAPID DATA PROCESSING

by

TADEUSZ WOJTOWICZ M.Sc., D.I.C.

Thesis submitted for the degree of

Doctor of Philosophy.

1971

CONTENTS

	Page
ABSTRACT	4
CHAPTER I - Optical Processing	5
I.1 Mathematical Analysis of Processing Operations	6
I.2 Fourier Transform Property of Lenses	12
I.3 Coherent Optical Processing Systems	15
I.4 Binary Spatial Filter	19
I.5 Amplitude Spatial Filter, Phase Spatial Filter	25
I.6 Complex Spatial Filters	28
I.7 Computer Generated Filters	32
CHAPTER II - Behaviour of Emulsions in Coherent Light	39
II.1 Characteristic Curve	40
II.2 Holographic Recording and Linearity	43
II.3 Non linearity	49
II.4 Dynamic Range and Noise	50
II.5 Diffraction Efficiency	52
II.6 Optical Transfer Function	55
II.7 Phase Effects	60
CHAPTER III - The Mach-Zehnder Interferometer	69
III.1 Coherence	69
III.2 Fringe Localisation	77
III.3 The Interferometer	82
CHAPTER IV - The Coherent Processor	95
IV.1 Diffraction Properties of a Circular Aperture	95
IV.2 The Coherent Processor	99

Page

CHAPTER V - The Photographic Emulsion 106

V.1 Physical Characteristics of the Emulsion 109

V.2 Emulsion Characteristics 113

V.3 Phase Measurements 124

V.4 Relief Images 136

V.5 Diffraction Efficiency 152

V.6 Effects of Base Phase Variations on Diffraction Plane 156

V.7 Scatter 187

V.8 Modulation Transfer Function 190

CHAPTER VI - Liquid Gates 195

VI.1 Design of the Liquid Gate 196

VI.2 Effect of Liquid Gate 201

CHAPTER VII - Conclusions 210

ACKNOWLEDGEMENTS 214

REFERENCES 215

APPENDIX 219

ABSTRACT

As the title suggests, the object of the project was to study some of the properties of photographic emulsions with particular reference to using rapid access emulsions in coherent optical processing systems. The first chapter describes some of the operations which can be performed using such systems. It is basically a chapter on optical data processing techniques. The second chapter discusses generally some of the properties of photographic emulsions and some basic equations are derived which are later investigated experimentally.

The third and fourth chapters are used to describe the apparatus which has been constructed in order to perform the experiments. Some of the well known background theory is also given so as to make the chapters more complete and self-contained.

Details of results and experimental work are presented in chapters five and six. The fifth chapter describes some of the properties of this type of emulsion and an attempt is made to correlate some of the results with equations which have been derived theoretically. Results are obtained showing how the flatness of the emulsion varies with area, and the influence of the flatness on the diffraction plane. The sixth chapter shows how the performance of the emulsion can be improved by using a liquid gate.

Finally, the last chapter gives a brief summary of the more important results obtained and some conclusions are reached as to the possibility of using this type of emulsion in automatic inspection systems utilising coherent light.

CHAPTER I - OPTICAL PROCESSING

When a signal is recorded photographically it can be inspected in several ways. In the present chapter we will be dealing with some of the operations which can be performed using lens systems and filters. These operations can then be used in order to process the signals to obtain the required information.

Various electronic data processing techniques have been known and used for many years. The input to such a processor takes the form of voltage variations, these signals are then processed by electrical circuits and the output can simply be read off. Such an electrical system is limited due to the fact that the input is only a one dimensional voltage display. The great advantage of optical processing systems is that the input takes the form of a two dimensional display of a photographic transparency, or some other similar suitable medium, thus enabling a great deal of information to be processed simultaneously.

A coherent optical signal processor is a device utilising coherent light sources and the diffraction phenomena of lenses to perform certain operations on the input signals. Operations which may be performed include filtering, multiplication, spectral analysis and combinations of these to do auto-correlations, cross-correlations and convolutions. An optical processor consists of a coherent light source, collimating optics, an input signal which is often in the form of a photographic transparency, additional optics to perform the operations required and image the signal onto a reference plane, and detection by either electro-optical means or an emulsion.

If a parallel beam of light is incident onto an input transparency placed a focal distance, F, from a converging spherical lens, then at a focal distance on the output side of the lens a pattern is displayed which is the two dimensional Fourier transform of the input signal (see Fig. 1). For the case mentioned the light distribution in the focal plane is a function of the spatial frequency content of the input signal. The Fourier transform can therefore be used to analyse a non-periodic function into its component spatial frequencies. The original input signal is uniquely described by its transform. If a second lens is then placed a focal distance from the transform plane an inverse transform is performed and the original image is then produced. The essence of optical processing is to use such a system in conjunction with a set of filters to perform certain operations on the signal.

I.1 Mathematical Analysis of Processing Operations

The following analysis can be found in most books on Fourier transforms, the treatment given below closely follows that given by Stroke.[9, 57, 75]

(a) Fourier Transform

The Fourier transform F(u) of the function f(x) is defined as

$$F(u) = \int_{-\infty}^{\infty} f(x)\, e^{-i 2\pi u x}\, dx \qquad (1.1)$$

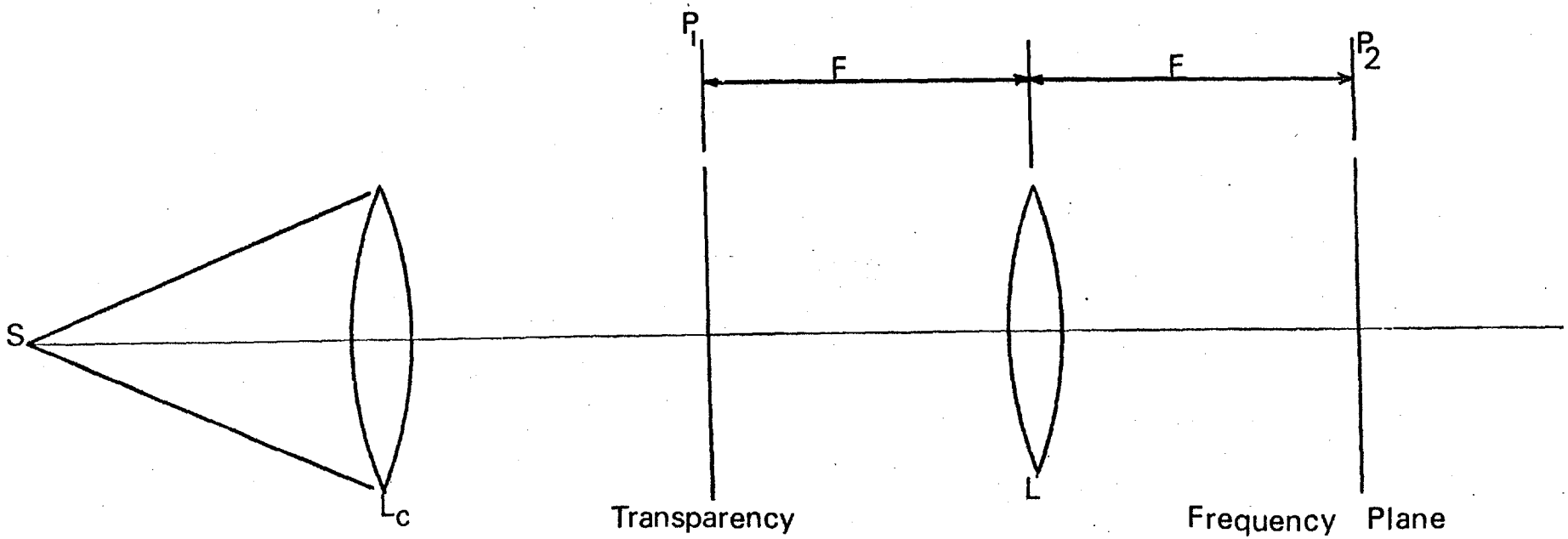


Fig. 1 Coherent Processing System

which is often written as

$$T[f(x)] = F(u) \tag{1.2}$$

where T represents the Fourier transform.

Similarly the inverse Fourier transform may be defined for F(u) as

$$f(x) = \int_{-\infty}^{\infty} F(u)\, e^{+i2\pi ux} \cdot du \tag{1.3}$$

or

$$T^{-1}[F(u)] = f(x) \tag{1.4}$$

The above equations are valid for functions of two or more variables. For function of two variables 1.2 and 1.4 become

$$T[f(x,y)] = F(u,v) \tag{1.5}$$

and

$$T^{-1}[F(u,v)] = f(x,y) \tag{1.6}$$

where

$$f(x,y) = \iint_{-\infty}^{\infty} F(u,v)\, e^{i2\pi(ux+vy)} \cdot du\, dv \tag{1.7}$$

If the function to be transformed is an accurate specified description of a physical quantity then the existence of the transform can be ignored. Two conditions often imposed on function f(x) to have a Fourier transform are that, x is a real variable ranging between $\pm\infty$ and the convergence condition that f(x) may be real or

complex but such that

$$\int_{-\infty}^{\infty} |f(x)|\,dx$$

exists.

Generalised functions such as the delta function $\delta(x)$ and the step function $H(x)$ do not fit this convergence condition. They can however be included in the Fourier transform theory since they are defined by a limiting process involving sequences of ordinary functions which do have transforms (Lighthill,[57]).

The following are some properties of Fourier transforms often useful in optics; they are given without proof.

$$\mathcal{T}[a_1 f_1(x)] + \mathcal{T}[a_2 f_2(x)] = a_1 F_1(u) + a_2 F_2(u) \tag{1.8}$$

where a_1 and a_2 are any complex constants;

$$\mathcal{T}[f(ax)] = \frac{1}{|a|} F\left(\frac{u}{a}\right) \tag{1.9}$$

$$\mathcal{T}[f^*(-x)] = F^*(u) \tag{1.10}$$

where $f^*(x)$ describes the complex conjugate of the function $f(x)$

$$\mathcal{T}[f(x \pm x_0)] = e^{\pm i 2\pi u x_0} \cdot F(u) \tag{1.11}$$

(b) Convolution

The convolution integral $f(x)$ is by definition obtained from two functions $f_1(x)$ and $f_2(x)$ according to

$$f(x) = \int_{-\infty}^{\infty} f_1(u) f_2(x-u)\,du \tag{1.12}$$

which can be written as $f = f_1 * f_2$, where * denotes convolution.

It can be shown that

$$F(u) = F_1(u)F_2(u) \tag{1.13}$$

The function f(x) is obtained by successive evaluation of the integral of the product of the two functions $f_1(x)$, $f_2(x)$ as the two functions are shifted with respect to each other. One value of the integral is obtained for each shift position x of $f_2(x)$ with respect to $f_1(x)$.

The convolution theorem is given by

$$\mathcal{T}[f_1 * f_2] = F_1 \cdot F_2 \tag{1.14}$$

and

$$\mathcal{T}^{-1}[F_1 \cdot F_2] = f_1 * f_2 \tag{1.15}$$

Equation 1.14 states that the Fourier transform of the convolution of two functions is given by the product of the transform of the two functions. Equation 1.15 states that the inverse Fourier transform of two functions is the convolution of the two functions.

(c) Correlation Functions

The autocorrelation function, $C_a(\tau)$, of F(t) is given by

$$C_a(\tau) = \int_{-\infty}^{\infty} F(t)^* F(t+\tau)\, dt \tag{1.16}$$

For function F(t) = constant we find that F(t) at $t + \tau$ is the same as at t, i.e. it is correlated. For a periodic function if we find that a shift of τ_0 each time produces the same value then the values at t and $t + m\tau_0$ are perfectly correlated. If we move each time by amount τ (which is any value) we find that the function

$C_a(\tau)$ has different values, i.e. the function is uncorrelated. If in an autocorrelation we see peaks at τ_0 intervals, the function is said to be correlated with period τ_0.

The definition of a cross-correlation function is similarly given by

$$C_c(\tau) = \int_{-\infty}^{\infty} F(t)^* G(t+\tau)\, dt \qquad (1.17)$$

It can be shown that the Fourier transform of the correlation function

$$c(v) = f^*(v)\, g(v) \qquad (1.18)$$

which is known as the Wiener Khinchin theorem. For the case of the autocorrelation function equation 1.18 reduces to

$$c(v) = |f(v)|^2$$

That is the spectrum of the autocorrelation function is the distribution of power in the spectrum (power spectrum) or the autocorrelation function of f is the inverse transform of the power spectrum of f. The output of a convolution operation is of the form of the original signal, that of a correlation operation is in frequency terms. Correlation requires the use of one lens and some scanning system whereas a convolution requires two lenses.

Basically what is involved in optical data processing is correlating unknown characters, the input, with an array of known characters, the filter, which are required. The output is in the form of some d.c. noise level with a series of correlation peaks corresponding to the required pattern. Difficulties arise when recognition is required between two similar patterns as both would produce a correlation function, the system is then required to find the degree of correlation.

I.2 Fourier Transform Property of Lenses

From Fig. 2 the complex amplitude distribution of light in the focal sphere is given by Huygens' principle.

$$u_p = \frac{i}{\lambda}\iint_A u_o \frac{e^{-ikR}}{R} \cdot dA$$

if we have a perfectly corrected spherical wavefront

$$u_o = f(x,y)\frac{e^{ikR_o}}{R_o}$$

therefore $$u_p = \frac{i}{\lambda}\iint_A f(x,y)\frac{e^{ik(R_o-R)}}{RR_o} \cdot dx\,dy$$

now $$R^2 = (x-\xi)^2 + (y-\eta)^2 + [z-(R_o-\Gamma)]^2 \qquad (1.19)$$

and $$R_o^2 = \xi^2 + \eta^2 + (R_o+\Gamma)^2 \qquad (1.20)$$

using equation 1.19 and 1.20 and assuming R_o-R is small it can be shown that

$$R_o - R = \frac{x\xi + y\eta}{R_o}$$

then $$u_p = \frac{i}{\lambda R_o^2}\iint_A f(x,y)\, e^{ik\left(\frac{x\xi+y\eta}{R_o}\right)}\, dx\,dy \qquad (1.21)$$

applying 1.21 to lens system of Fig. 3 the entrance pupil is of the form $x^2 + y^2 = 1$

let P be P(X, Y), $x = X/h$ and Y/h

we have $$\frac{n'\sin\alpha'}{\lambda}\eta' = \frac{n\sin\alpha}{\lambda}\eta \qquad (1.22)$$

using reduced co-ordinates

$$u = \frac{n\sin\alpha}{\lambda}\xi \quad , \quad u' = \frac{n'\sin\alpha'}{\lambda}\xi'$$

$$v = \frac{n\sin\alpha}{\lambda}\eta \quad , \quad v' = \frac{n'\sin\alpha'}{\lambda}\eta'$$

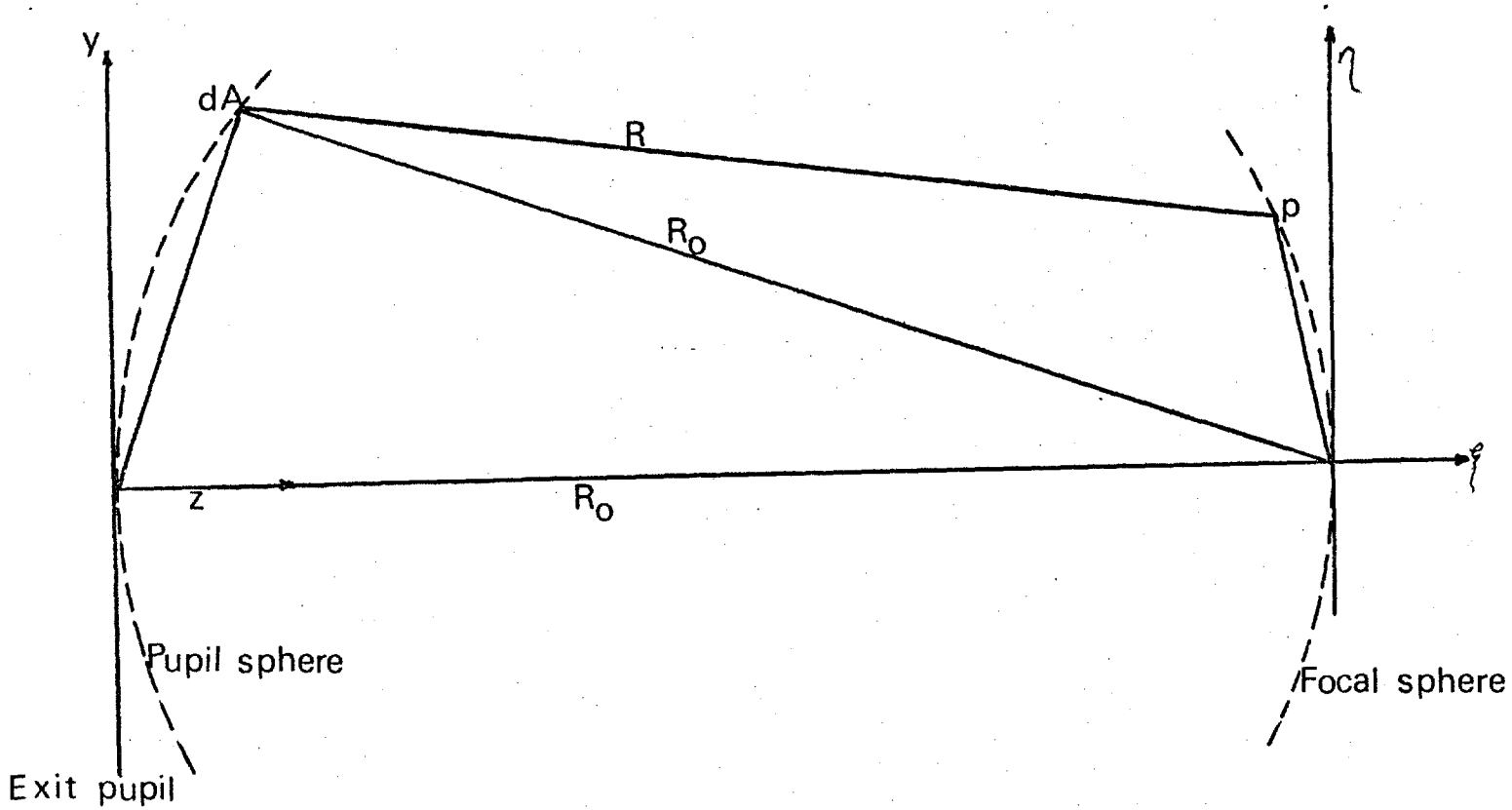


Fig 2 Pupil – Focal sphere relationship

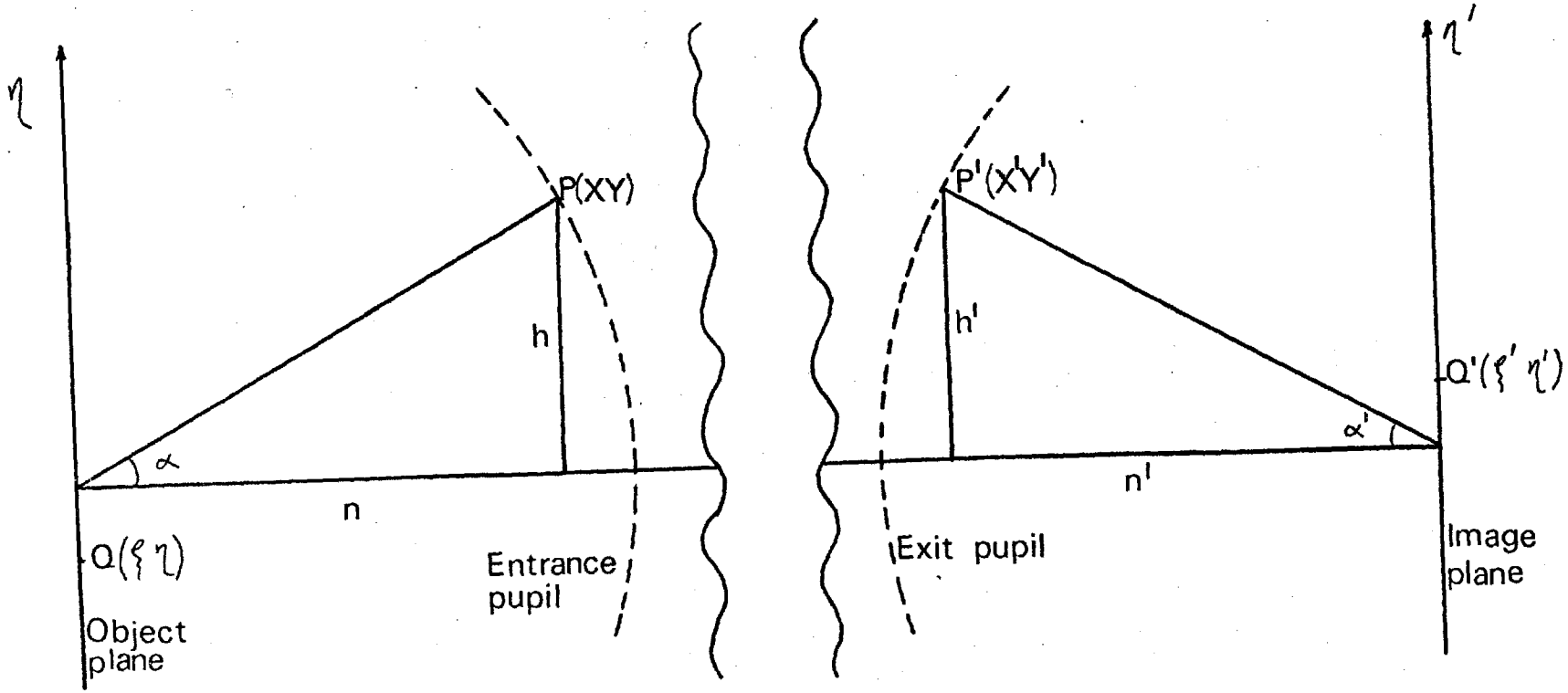


Fig 3 Fourier Transform

for exit pupil $x' = {}^{X'}/_{h}$ and $y' = {}^{Y'}/_{h}$ from 1.22 $v' = v$ and $u' = u$ using 1.21 the complex amplitude in the image sphere is given by

$$F(u,v) = K \iint_A f(x,y)\, e^{i2\pi(ux+vy)}\, dx\,dy \qquad (1.23)$$

Thus we see the Fourier transform property of lenses. Equation 1.23 is the exact relationship, in coherent optical systems however the lenses are usually being used at infinite conjugates and for such cases the pupil spheres become planes and the image sphere becomes the image plane. The planes of interest are at the focal distance, F, on each side of the lens and it is to these planes that we apply our Fourier transform equations.

I.3 Coherent Optical Processing Systems

If a transparency at plane P_1, (Fig. 1) is illuminated by a monochromatic plane wave of light, then at plane P_2 the distribution of light is given by equation 1.23. The variables u, v, are spatial frequencies, and x, y, are lengths and the variables u, v are related to distances ξ , η in plane P_2 by $u = \frac{2\pi\xi}{\lambda F}$ and $v = \frac{2\pi\eta}{\lambda F}$. Hence we see that the complex light distribution in plane P_2, given by the Fourier transform of the input plane P_1, is the frequency content of the input signal. The upper limit of the spatial frequency which can be displayed in P_2 is dependent upon lens L_1. If we place the input signal closer to L_1 the Fourier transform relationship is still valid, except for a phase factor, and the bandwidth is increased. If we were to use spherical waves rather than parallel light waves, the frequency plane would be displayed at the point of focus of the spherical wave. Fig. 1 shows a typical coherent optical processing system. S is monochromatic point source,

L_c collimating lens, P_1 the input plane function described by f(x, y), P_2 the frequency plane, $F(\xi,\eta)$. This type of system is known as a two dimensional spectrum analyser. If a cylindrical lens is used in combination with spherical lens L_1, the cylindrical lens has a focusing effect in one dimension only and hence we are able to convert a two dimensional system into a multichannel astigmatic optical system.(17)

If we now add another lens L_2 to the system as shown in Fig. 4 the lens L_2 performs an inverse Fourier transform on the plane P_2. The light distribution in plane P_3 is then the same as in plane P_1 but with reversed co-ordinates. If we now wish to modify the original signal in some way we can do this by modifying the amplitude and phase in the frequency plane P_2. A filter is inserted at plane P_2 which transmits the spectra with the required attenuation. This is known as a spatial frequency filtering operation. Suppose the complex amplitude transmittance of the filter is given by H(u, v) then the modified light distribution at P_2 is given by R(u, v) = F(u, v) H(u,v). The distribution of light in plane P_3 is given by the Fourier transform of R(u,v)

$$r(x,y) = \iint_{-\infty}^{\infty} F(u,v)\,H(u,v)\,e^{i2\pi(ux+vy)}\,du\,dv \qquad (1.24a)$$

or by using the convolution theorem and the impulse response h(x,y) of the filter 1.24a becomes

$$r(x,y) = \iint_{-\infty}^{\infty} f(x',y')\,h(x-x',y-y')\,dx'dy' \qquad (1.24b)$$

provided all the spectra of the input signal are transmitted by lens L_2.

We have already seen in section I.1(c) that we can perform

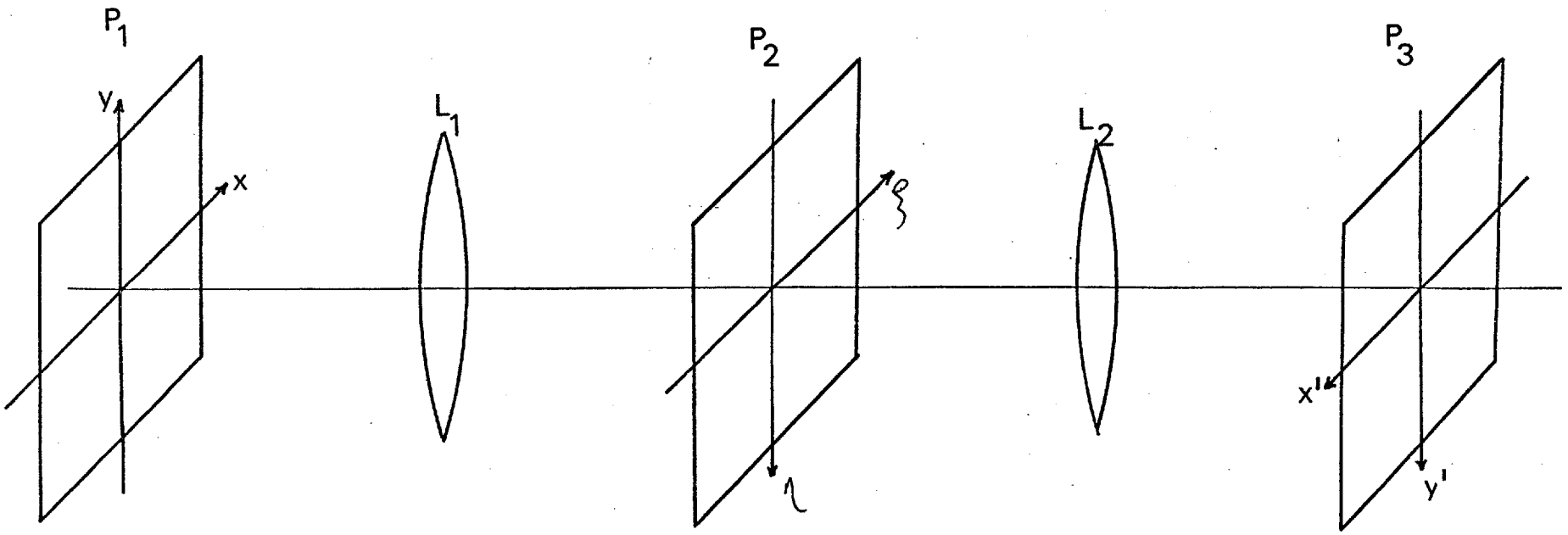


Fig 4 Fourier Transform System

filtering operations in the spatial domain. If a transparency s(x,y) is placed in this domain then the signal in the frequency plane is given by

$$r(u,v) = \iint_{-\infty}^{\infty} s(x,y) f(x,y) e^{-i 2\pi(ux+vy)} \cdot dx\,dy \qquad (1.25)$$

which is of the same form as equations 1.24. The two methods produce the same results but display the information differently.

The filter G(u,v) can be a complex function having amplitude response $|G(u,v)|$ and phase response $\exp[i\phi(u,v)]$

$$G(u,v) = |G(u,v)|\, e^{i\phi(u,v)}$$

We also know that $|G(u,v)| \leq 1$ since it can only attenuate rather than amplify the light distribution, hence the possible values of a complex filter must lie on or within the circumference of a unit circle centred about the origin of the complex plane.(17, 63)

Just as for the spectrum analyser we can modify the systems performing the above filtering operations from a two dimensional to a multichannel one dimensional system by using a cylindrical lens. The filter has the appropriate frequency response for each channel. The one dimensional convolution becomes

$$r(x,y) = \int_{-\infty}^{\infty} f(x',y)\, h(x-x',y)\, dx' \qquad (1.26)$$

with a similar expression for the correlation operation. In order to perform these operations we must first of all produce the required filter. The filters can be made either by photographic recording or generated from data calculated by computer. The various different types of filter are dealt with in the following sections.

I.4 Binary Spatial Filter

This is the simplest form of filter, it is made up of opaque and clear areas. A stopband in the optical filter is the opaque portion and a passband is simply a transparent area. These and other types of filter have been dealt with by several authors.[7,11,45,63,65,67,81]

The simplest form of binary filter can be made up by having a central circular stop surrounded by an opaque screen. This has the effect of only allowing the low frequency content to pass through, when placed at the centre of the frequency plane; it is known as a low pass filter. The reverse of this is the high pass filter which consists of a central opaque spot with transparent surroundings. For the case of one dimensional spatial filtering the above would consist of a clear or opaque slit. More complicated binary filters are made up from combinations of such simple filters thus transmitting only the required frequency or frequencies. The performance of the compound filter can be derived from the performance of the individual simple filters. It is therefore necessary to know the performance of a simple filter both as a function of its size and also its position in the spatial frequency plane. The modified light distribution in the image, when using such a filter, can be found by either performing a Fourier transform of the modified frequency spectrum or by the convolution of the spread function, associated with the filter, with the perfect image. The Fourier transform method is simplest to use and will be used in the following analysis.

Consider a half plane object positioned at the origin of plane P_1 (Fig. 4). This may be represented by a function H(x) where

$$H(x) = 1 \text{ for } x > 0$$
$$H(x) = 0 \text{ for } x < 0$$

A coherent monochromatic plane wave incident onto this half plane will produce a distribution of complex amplitude in the frequency plane P_2 given by the Fourier transform of H(x)

$$h(u,v) = \tfrac{1}{2}\left[\delta(u) - \frac{i}{\pi u}\right]\delta(v) \qquad (1.27)$$

where δ is the Dirac delta function which only has a value at u = 0 and

$$\int_A \delta(u)\,du \quad = 1 \text{ for A including } u = 0$$
$$= 0 \text{ for A not including } u = 0.$$

The simple spatial filter is placed in plane P_2. In one dimension this is a slit placed at some position in P_2, in two dimensions it is a circular aperture. Let us consider the one dimensional case of an asymmetrical slit.(7) The transmission function of the filter t(u) is

$$t(u) = 1 \text{ for } -u_1 < u < u_2$$
$$= 0 \text{ for } -u_1 > u > u_2$$

where u_1 and u_2 are defined in Fig. 5. For the two dimensional circular aperture $t(u,v) = 1$, $u^2 + v^2 < 1$ and 0 outside.

The amplitude distribution in the image plane P_3 is found from the inverse Fourier transform of the complex amplitude transmission of the frequency plane P_2. This is given by the product of the object frequency spectrum and the filter transparency function t(u)

$$A(x') = \int_{-u_1}^{u_2} h(u)\,t(u)\,e^{i2\pi u x'}\,.\,du \qquad (1.28)$$

where h(u) is the one dimensional form of 1.27.

As shown in Fig. 5 the limits of the integration of 1.28 can be split up into limits between $-u_1$ and u_1, and u_1 and u_2. Equation

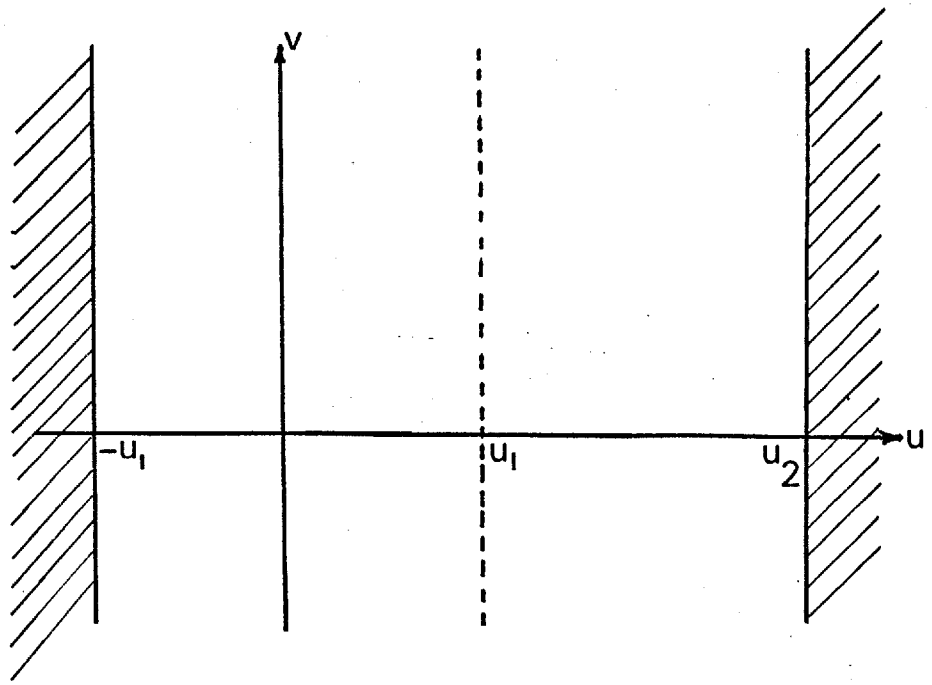


Fig 5 Slit Filter

1.28 now becomes

$$A(x') = \int_{-u_1}^{u_1} h(u)\, e^{i2\pi u x'}\, du + \int_{u_1}^{u_2} h(u)\, e^{i2\pi u x'}\, du \qquad (1.29)$$

substituting for h(u) and considering the first integral only.

$$A_1(x') = \int_{-u_1}^{u_1} \tfrac{1}{2}\left[\delta(u) - \tfrac{i}{\pi u}\right] e^{i2\pi u x'}\, du$$

$$= \tfrac{1}{2} - \tfrac{i}{2\pi} \int_{-u_1}^{u_1} \tfrac{1}{u}\, e^{i2\pi u x'}\, du.$$

Splitting up the exponential into sine and cosine and realising the integral involving cosine reduces to zero we have

$$A_1(x') = \tfrac{1}{2} - \tfrac{i}{2\pi} \int_{-u_1}^{u_1} \tfrac{i}{u} \sin 2\pi u x'\, du$$

now $\int_0^z \frac{1}{t} \sin t.dt = S_i(z)$

therefore $\int_0^z \frac{1}{t} \sin(at).dt = S_i(az)$

therefore $A_1(x') = \tfrac{1}{2} + \tfrac{1}{\pi} S_i(2\pi u_1 x') \qquad (1.30)$

Considering the second integral of 1.29

$$A_2(x') = \int_{u_1}^{u_2} \tfrac{1}{2}\left[\delta(u) - \tfrac{i}{\pi u}\right] e^{i2\pi u x'}\, du$$

The value of $\delta(u)$ is only non zero at u = 0 which is not within the limits of the integral, therefore the integral of the delta function reduces to zero and

$$A_2(x') = -\tfrac{i}{2\pi}\left[C_i(2\pi u_2 x') - C_i(2\pi u_1 x') + i\left\{S_i(2\pi u_2 x') - S_i(2\pi u_1 x')\right\}\right] \qquad (1.31)$$

where $\int_0^z \frac{1}{t} \cos t.dt = C_i(z)$

let $U_1 = 2\pi u_1 x'$, $U_2 = 2\pi u_2 x'$

The total distribution of amplitude in plane P_3 is given by 1.30 and 1.31.

$$A(x') = \tfrac{1}{2\pi}\{\pi + [S_i(u_1) + S_i(u_2)] - i[C_i(u_2) - C_i(u_1)]\} \tag{1.32}$$

The intensity distribution in P_3 is given by

$$I(x') = A(x')\,A^*(x')$$

$$I(x') = \tfrac{1}{4\pi^2}\{[\pi + S_i(u_1) + S_i(u_2)]^2 + [C_i(u_2) - C_i(u_1)]^2\} \tag{1.33}$$

This gives the distribution of intensity of the image at all points except x' = 0 as the cosine integral has a value of minus infinity for zero argument. The value of A(x'=0) is found from

$$A(x'=0) = \tfrac{1}{2}\int_{-u_1}^{u_2}\{\delta(u) - i/\pi u\}\,du$$

$$= \tfrac{1}{2}\int_{-u_1}^{u_1}\delta(u)\,du - \tfrac{i}{2\pi}\int_{-u_1}^{u_1}\tfrac{1}{u}\,du - \tfrac{i}{2\pi}\int_{u_1}^{u_2}\tfrac{1}{u}\,du$$

$$= \tfrac{1}{2} - \tfrac{i}{2\pi}\log_e\left(\tfrac{u_2}{u_1}\right) \tag{1.34}$$

From equations 1.33 and 1.34 we are able to find the distribution of intensity at all points in the image for an asymmetric centred slit filter. For a large slit displaced from the centre of the frequency spectrum such that the filter transmission function t(u) has a value of zero for $\pm u_1$, the distributions of amplitude and intensity in the image plane become

$$A(x') = -\tfrac{i}{2\pi}[C_i(u_2) - C_i(u_1) + i\{S_i(u_2) - S_i(u_1)\}] \tag{1.35}$$

$$I(x') = \tfrac{1}{4}\pi^2\left[\{C_i(u_2)-C_i(u_1)\}^2 + \{S_i(u_2)-S_i(u_1)\}^2\right] \qquad (1.36)$$

$$I(x'=0) \doteq \tfrac{1}{4}\pi^2 \log_e(u_2/u_1) \qquad (1.37)$$

Using the above equations it is possible to evaluate the distribution of light intensity in the image for various configurations of compound filters[7] when the object is in the form of a half plane.

One of the earliest examples of binary filtering was described by Abbe in his theory on the image formation in a microscope. In his experiments Abbe saw a Fraunhofer spectrum due to a coherently illuminated grating. For a pure sine wave grating the spectrum would be the d.c. component plus one spectrum on either side. It is the mutual interference of these spectra which gives rise to the final image. If the aperture of the system was stopped down so that only the d.c. component was allowed through, then the image would be seen to be of uniform illumination with no structure. If the spectrum of a grating consists of a d.c. component and a series of harmonics, then all the harmonics must pass through the system for a faithful image to be observed. If an annular stop is inserted in the frequency plane then the frequency content of the image can be drastically altered.

Binary filters are useful for detecting certain required periodic signals in the presence of noise and other signals. The filter is constructed to allow through only the spectra of the required signal. Harmonics of other spectra falling at the same point in the frequency plane would also be transmitted but their signals would be much

fainter and so the signal to noise ratio could be greatly increased. By using such a filter to eliminate the low frequencies the contrast of a photograph can be increased.

I.5 Amplitude Spatial Filter, Phase Spatial Filter

Unlike the binary spatial filter which only has transmission values of zero and one, the amplitude filter can have any value between zero and one for any frequency. It is sometimes used rather than a binary filter where the sharp edge of the binary filter could produce its own diffraction image and interfere with the required image. The smooth transmission of an amplitude filter eliminates this diffraction effect.

Amplitude filters are usually recorded on photographic film, they are not often used on their own but in conjunction with a phase filter. In fact it would often prove to be difficult, if not impossible to record an amplitude filter on film without having any phase effect present. When plotted in the complex plane the amplitude filter has only real values.

Phase filters have no effect on the amplitude but either advance or retard the phase. In the complex plane they have values between plus and minus i on the imaginary axis. Uses of such filters are limited due to the difficulties encountered in their production. They can be made by using the relief image properties of photographic film (see following chapter), by deposition of a film onto a substrate, by the figuring of a piece of glass, they can also be made holographically as described in a later section, I.6. It is therefore seen that it is difficult to make a phase filter with continuous phase variations. Phase filters are normally made to

advance or retard the phase of the light by one half wavelength.

The main use of phase filters is to change a phase variation into an intensity variation. Suppose we have a phase object defined by

$$A(x,y) = e^{i\,\Phi(x,y)}$$

for $|\Phi(x,y)|$ small compared to unity

$$A(x,y) = 1 + i\,\Phi(x,y) \tag{1.38}$$

Referring to Fig. 4 the object given by 1.38 placed at the plane P_1 would produce at P_2 the Fourier transform of A(x,y). Therefore at P_2 we have

$$a(u,v) = \delta(u,v) + i\,\phi(u,v) \tag{1.39}$$

where $\delta(u,v)$ is the Dirac delta function and $\phi(u,v)$ is the Fourier transform of $\Phi(x,y)$. If in this plane we place a phase filter which has the transmission characteristics

$$t(u,v) = \pm i \text{ at } u = 0,\ v = 0$$
$$= 1 \text{ elsewhere}$$

then we have advanced or retarded the delta function, since it occurs at the point u = v = 0, but left the rest unchanged. The function emerging from plane P_2 is now

$$a'(u,v) = \pm i\,\delta(u,v) + i\,\phi(u,v)$$

which is imaged by the lens L_2 to give

$$A(x',y') = \pm i + i\,\Phi(x,y) \tag{1.40}$$

The intensity in the image is obtained from 1.40 by multiplying it with its complex conjugate.

$$I(x',y') = 1 + 2\,\phi(x,y) + [\phi(x,y)]^2 \qquad (1.41)$$

The last term of 1.41 can be ignored as $\phi(x,y)$ is assumed small. Therefore we have produced a variation of intensity of the image which follows the variations of phase of the original object. We have changed a phase variation, which cannot be seen, into a visible intensity variation.

Phase filters are usually used in conjunction with amplitude filters, so that such a composite filter can assume all values within the complex plane of a unit circle. If such a combination was used for the previous example, 1.41 would become

$$I(x',y') = T^2 + 2T\,\phi(x,y) \qquad (1.42)$$

where T is the amplitude attenuation. Such filters are used in phase contrast techniques for viewing pure phase objects. If the +i filter is used we have bright field contrast, the -i filter produces dark field phase contrast.

Tsujiuchi[80] demonstrated the use of amplitude and phase filters in correcting aberrated images. In one of his methods he uses a composite filter in the exit pupil of a lens imaging system to correct for the wave aberration. If the wavefront is given by $w(x,y)$, then the correcting filter would be of the form of $H(x,y) = \exp[-ikw(x,y)]$ which is a pure phase filter. Owing to the difficulty in making such a filter, except in special cases, he uses an amplitude and phase filter combination to minimise the aberrations whilst using a simple phase filter function.

I.6 Complex Spatial Filters

So far we have been dealing with filters which change either the phase or the amplitude of a signal. A more useful type of filter is one which can do both operations, taking any value within a unit circle in the complex plane; this is a complex filter. Such a filter could be used to detect a known signal in the presence of random noise, where the noise could be the rest of the scene of which the required signal is only a part. Suppose s(x,y) is the required known signal, and n(x,y) is the random noise. We need to find the optimum filter, h(x,y) which will operate on f(x,y) = s(x,y) + n(x,y) to maximise the ratio of the peak signal energy to the mean square noise energy. It can be shown[81] that the required filter response is

$$H(u,v) = kS^*(u,v)/N(u,v) \qquad (1.43)$$

Where S*(u,v) is the complex conjugate of the Fourier transform of the signal, N(u,v) is the noise spectrum, and k is a constant chosen so that $|H(u,v)| \leq 1$ for passive filters. For the case of the matched filter N(u,v) is uniform throughout the frequency plane. This is often not true for optical filters.

Equation 1.43 gives us the characteristics of the optimum filter, we now have to find some method of producing this function. N(u,v) is always a non-negative function which can be recorded as photographic film by making the exposure proportional to N(u,v) and having a gamma of 2.

Problems arise in realising S*(u,v). Since this could be a complex quantity we have to record both amplitude and phase. If we simply use a spherical lens to produce S(u,v) by performing a Fourier

transform of s(x,y), we cannot record this because physical detectors can only record the intensity of S(u,v) and not its phase. Vander Lugt[65] describes a system whereby the optimum filter can be realised. In his system he uses a modified Mach-Zehnder interferometer, (Fig. 6). The required signal s(x,y) and a spherical lens are placed in one arm of the interferometer. The Fourier transform of s(x,y) is then displayed at the back focal plane of the lens, the observed output is a combination of this and the spectrum due to the reference beam R(u,v). The output at the back focal plane is given by

$$G(u,v) = |R(u,v) + S(u,v)|^2$$

$$= |R(u,v)|^2 + |S(u,v)|^2 + R^*(u,v)S(u,v) + R(u,v)S^*(u,v) \qquad (1.44)$$

where both R(u,v), S(u,v) are complex. $R(u,v) = |R(u,v)|e^{i\phi(u,v)}$ and $S(u,v) = |S(u,v)|e^{i\theta(u,v)}$. We can write 1.44 now as

$$G(u,v) = |R(u,v)|^2 + |S(u,v)|^2 + 2\,\text{Real pt.}[R(u,v)S^*(u,v)]$$

$$= |R(u,v)|^2 + |S(u,v)|^2 + 2|R(u,v)||S(u,v)|\cos[\phi(u,v) - \theta(u,v)] \qquad (1.45)$$

By recording G(u,v) on photographic film we also record the phase information in S*(u,v). The film is exposed so that its transmission is proportional to G(u,v) and combined with the film on which $^{1}/_{N}(u,v)$ was recorded. The transmission of the two films is then given by

$$\frac{G(u,v)}{N(u,v)} = \frac{|R(u,v)|^2 + |S(u,v)|^2}{N(u,v)} + \frac{R^*(u,v)S(u,v)}{N(u,v)} + \frac{R(u,v)S^*(u,v)}{N(u,v)} \qquad (1.46)$$

If we make |R(u,v)| a constant k and its phase $\phi(u,v)$ linear in u and v, $\phi(u,v) = au + bv$, then the third term of 1.46 becomes

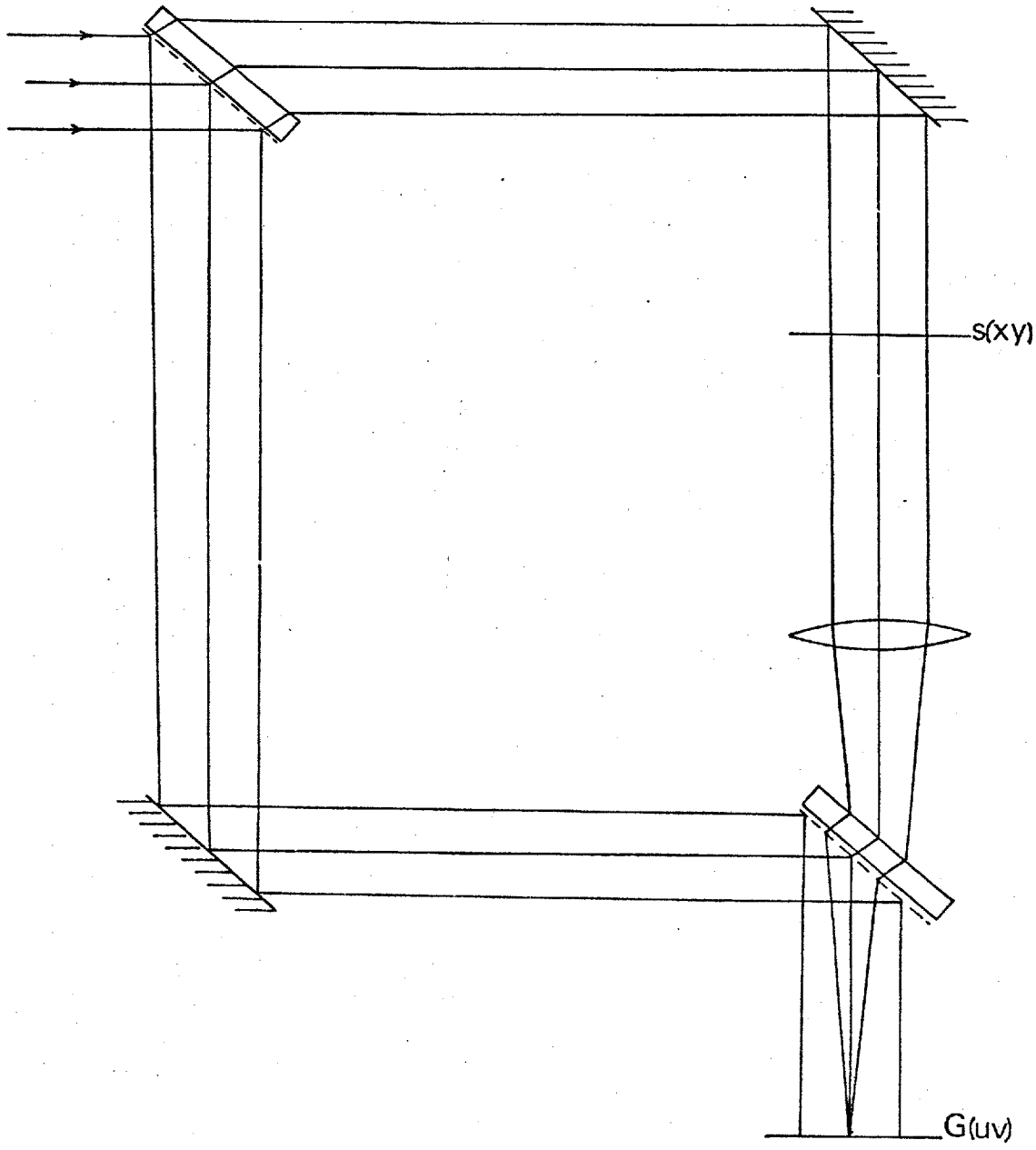


Fig 6 Formation of Complex Spatial Filter

$$\frac{k S^*(u,v)}{N(u,v)} e^{i(au+bv)} = H(u,v) e^{i(au+bv)}$$

which is the desired filter function as given by 1.43 multiplied by a linear phase factor. If we place this filter into the frequency plane of the coherent optical processor (Fig. 4), the output in plane P_3 is given by the inverse Fourier transform. The original signal was $f(x,y) = s(x,y) + n(x,y)$ at P_1 producing $F(u,v)$ at P_2. The combination of this and the filter would produce at plane P_3

$$f(x',y') = \iint_{-\infty}^{\infty} F(u,v) H(u,v) e^{i(ux'+vy')} \, du\,dv \qquad (1.47)$$

The effect of placing the filter given by 1.46 into plane P_2 would be for the lens L_2 to separate out the three terms of 1.46. The first term appears on axis, the second and third terms on opposite sides of the optic axis. The third term is given by

$$\iint_{-\infty}^{\infty} F(u,v) H(u,v) e^{i[(a+x')u+(b+y')v]} \, du\,dv \qquad (1.48)$$

which is the output as required by the optimum filter.
The centre of 1.48 is located at $x' = -a$, $y' = -b$ and can be picked up by a detector. To prevent overlap of the signals the values of a and b are made such that they are greater than the dimensions of the signal.

The use of these complex filters for signal detection has been demonstrated by several authors.[11, 45, 65] They have been able to detect a particular shape or letter from a signal containing many different shapes and letters. Both convolution and correlation techniques have been used. It is of interest to note that when using these correlation techniques the relative intensity of detection

is a function of the degree of correlation. Different spatial orientations of a particular shape give different relative correlation peaks. Also one often finds that two different but similar shapes or letters have some degree of correlation and can be confused in detection.

As an extension to using these filters for detecting a single signal, it is sometimes required to detect several signals simultaneously from a large data display. It is possible to record several signals on a single photographic film by using multiplexing techniques.(64) The solution however now becomes complicated because of the greater risk of finding correlation peaks due to two different signals thus confusing the output.

I.7 Computer Generated Filters

The production of filters to perform operations on signals is not always simple. In recent years people have been using computers to generate and plot synthetic holograms (filters). These types of filters have distinct advantages over conventional optically produced filters. Firstly, we no longer require to have the original object or signal since this can be defined mathematically. Secondly, the optical elements used to construct the filter can be computed as perfect, we no longer have to worry about degradations produced by lenses and filters or non-uniform illumination. Thirdly, the recording of the filter can be much more accurately controlled, the correct bias point of the film recording characteristics (transmittance-exposure see next chapter) can more easily be achieved and the effect of non-linearities can be compensated for. There are, however, two disadvantages; degradations of the image occur due

to the limited spatial resolution and the availability of the intensity levels which the plotting equipment can produce.

Three different approaches have been used for synthesising computer generated filters:

(a) by using the "kinoform" device in which only the phase is plotted[55]

(b) by computing the distribution of light intensity at the recording plane, adding a reference wave, then plotting the resultant intensity distribution on film using a grey scale plotter[56]

(c) by generating a Fourier transform of the desired image with no adding of a reference wave, then using the Fourier transform properties of lenses to reconstruct the image.[11, 29, 51, 60, 61]

In method (a) Lesem, Hirsch and Jordan computed the wavefront reflected by a diffusely reflecting object illuminated with coherent light. They plotted this onto film, then discarded the amplitude information by a bleaching process. The resultant thickness modulation of the film is proportional to the phase distribution of the calculated object-wave. The phase hologram when placed in a collimated beam of coherent light reconstructs the original object. Kermisch[41] shows analytically the effect of discarding the amplitude information on the reconstructed image.

In method (b), Lesem, Hirsch and Jordan used a version of Kirchhoff's diffraction integral to evaluate the holographic process. They used the Leith-Upatnieks hologram arrangement where a plane wave passes through a diffuser onto the object. The light scattered by the diffuser and object is recorded on a photographic emulsion. A reference beam incident at an angle to the emulsion is used to

modulate the scattered light.

The two dimensional object, which can be thought of as an array of resolution elements, is characterised by a transmission function T(x,y). A propagation function F(x-a,y-b,z) is the appropriate complex valued function which describes the propagation of light from a point on the object to a point on the image plane. The hologram can then be calculated from the object by mapping the elements of the object onto the resolution elements of the hologram using the appropriate function. This however requires a large number of computer machine operations. Lesem, Hirsch and Jordan found it to be much quicker by first computing the Fourier transform of T(x,y) and F(x-a,y-b,z) using the Cooley-Tukey[16] algorithm. The product of these two transforms can the be multiplied forming the transform of the convolution of T and F. The inverse transform of this convolution product is then the wavefront of the light scattered by the object at the hologram plane. To this is added the plane reference wave, and the intensity of illumination in the hologram plane is given by the squared modulus of this. The hologram can then be made by plotting the hologram intensity distribution onto film using a variable density plotter. Most plotters do not have sufficient spatial resolution to make usable holograms. In most cases photographic reduction techniques are used to improve the hologram plots. Waters[85] extended this technique to make holograms of three dimensional objects. He considered the three dimensional Fourier transform of an object in a volume centred about the focus of the transform lens. The relationship between the amplitude distribution in the hologram plane due to a total number of spherical waves coming from each point in the object is given by

$$U_s(u,v) = K\int_V u(x,y,z)\exp\left\{-\frac{ikz}{2}\left[\frac{z^2}{f^2}(u^2+v^2)+\frac{2}{f}(ux+vy)\right]\right\}dV \qquad (1.49)$$

where u(x,y,z) is the complex amplitude distribution of the object, the co-ordinate z is being the distance of the object from a point reference source which is at the focal distance f of the transform lens.

As the exponential of 1.49 is linear in object space co-ordinates (x,y,z), the integration can be performed for simple geometries such as straight lines, planes and simple curves for any orientation. A more complex object can then be thought of as a combination of these simple geometries, hence a three dimensional object can be transformed into a two dimensional plane by using the Fourier transform of 1.49. As before, a reference plane wave is added to this and the distribution of intensity in the hologram plane is evaluated.

For producing filters for data processing systems, it is sufficient to produce the Fourier transform of the desired signal, since it is this transform that will be used as a filter. For this reason method (c) is the most practical for making filters to be used in data processing systems.

We wish to produce a complex amplitude filter F(u,v) which is the Fourier transform of the desired object. We must therefore be able to reproduce both the amplitude and phase. Brown and Lohmann[11] devised a method to influence the phase which they called detour phase. Two rays leaving adjacent slits in a grating and going into the Mth diffraction order have a path difference of λ and detour phase of 2π. The grating equation gives us the path difference.

$$M\lambda = d(\sin\alpha_m - \sin\alpha_0) = L_m \qquad (1.50)$$

where d is the grating spacing, and α_0, α_m the incident and diffracted angles respectively. If we have some of the slits displaced by an amount P then the path differences at the displacement is given by

$$L'_m = (d + P)(\sin\alpha_m - \sin\alpha_0) \qquad (1.51)$$

thus forming a deformed diffracted wave. The wave emerging from the displaced portion of the grating is retarded by an amount $\Delta_m = L'_m - M\lambda$. From 1.50 and 1.51 we find that

$$L'_m - M\lambda = P(\sin\alpha_m - \sin\alpha_0) = \frac{PM\lambda}{d}$$

thus producing a phase delay or detour of phase of $\frac{2\pi PM}{d}$. Hence we are able to provide the required phase shift by having an irregular grating.

The amplitude of the required filter function can be realised on photographic film. However it is difficult to achieve the necessary grey scale to characterise the amplitude, instead a binary form of an array of tiny dots is used. If sufficiently small cells $(\Delta\delta)^2$ are chosen then the function F will be almost constant

$$F(u,v) \approx F(n\Delta\delta, m\Delta\delta) = F_{nm} = A_{nm}e^{i\alpha_{nm}}$$

where

$$|u - n\Delta\delta| \leq \Delta\delta/2 \;,\; |v - m\Delta\delta| \leq \Delta\delta/2 \;,\; n,m = 0 \pm 1 \pm 2 \ldots$$

We require that the complex amplitude F_{nm} will emerge from cell (n,m). This can be achieved by having the required binary structure inside each cell. If for example the cells are vertical slits then the

required filter function can be computed by calculating the required width and position of each vertical slit in each cell. The width of each slit would give the required amplitude and the position relative to the centre of the cell would give the required phase retardation. The desired filter function is produced by plotting the n.m cells which constitute an array of parallel slits. As the cells are binary in nature the hologram produced is of a binary structure and no grey tone scale is necessary. The original Fourier transform to obtain F(u,v) is performed using the Cooley-Tukey algorithm. The cells are drawn out at some suitable size, then reduced photographically. The sharpness of the edges due to the photographic reduction may be reduced but this does not affect the reconstructed image to a great extent.

These binary holograms are as efficient as ordinary thin emulsion holograms, the difference being that binary holograms are similar to square wave gratings, whereas normal holograms are sine wave gratings. The effect of the square wave grating is to produce harmonics which appear as higher order images. These images are as far away from the main image as the main image is from the zero order so that by avoiding overlap with the zero order one also avoids overlap with the higher orders. When using these square wave binary holograms $(\frac{4}{\pi})^2$ times more light is sent into the first diffracted order as compared to a perfect sine wave grating.

A further advantage of binary holograms arises from the fact that in some cases the photographic emulsion exhibits more grain noise at some intermediate transmittance value than at zero or one. Sometimes, therefore, these binary holograms are better as regards signal to noise ratio.

Lee[51] has recently described a new technique for generating holograms. He decomposes the original complex function into four quadrature components. The function $|F(u,v)| \exp[i\phi(u,v)]$ is expressed as

$$|F(u,v)|\exp[i\phi(u,v)] = F_1(u,v) - F_3(u,v) + i F_2(u,v) - i F_4(u,v)$$

each of the functions being real and non-negative. These four functions with the correct phases are used to produce F(u,v). The phase is obtained by using an appropriate "shift" parameter. Lee plots the holograms using 256 quantised levels to produce the required amplitude transmission. If more levels are required these can be produced by multiple exposures.

CHAPTER II - BEHAVIOUR OF EMULSIONS IN COHERENT LIGHT

The previous chapter dealt with the use of coherent illumination in data processing. The photographic process is a useful medium for recording data, other media include photoresists, photopolymers and photochromics. For the present case we are limiting ourselves to the use of photographic films, and the behaviour of these films in coherent optical systems.

In order to classify a photographic process one has to designate a set of parameters which would completely describe the process. An inspection of these parameters would then enable one to determine the best set of conditions and the best emulsion to use for a particular case, and also to know the limitations of the process under these conditions. As the photographic process is a rather complicated one, it can only be fully represented by a rather large number of mathematical functions. Some of these are listed below.

(a) Optical path differences
(b) Optical transfer function
(c) Diffraction efficiency
(d) Light scatter
(e) Linearity and dynamic range
(f) Speed
(g) Granularity
(h) Resolution
(i) Veiling glare
(j) Reciprocity failure
(k) Noise, signal to noise ratio
(l) Ageing characteristics

These parameters are influenced by certain processes and effects such as those listed below.

(1) Developers and developing conditions
(2) Bleaching process
(3) Atmospheric variations of temperature and humidity
(4) Gelatin shrinkage

(5) Emulsion strain

(6) Film damage and static charge

(7) Fluid gates

(8) Wavelength dependence

If a beam of coherent light is incident upon a clear circular aperture it will be diffracted. At some distance away a diffraction pattern is produced, which in the case of the circular aperture, is known as the Airy disc. As has already been seen in the previous chapter the distribution of light intensity in the diffraction pattern is the modulus of the square of the Fourier transform of the amplitude transmission of the aperture; for the circular aperture this is given by a Bessel function. If now a photographic emulsion is placed in the plane of the diffracting aperture, the diffraction pattern so formed is a representation of the emulsion. Any variation of complex transmission of the emulsion would create a change in the diffraction pattern. The function of photographic film in optical data processing is to produce the required variations in the complex transmission of the aperture, it is therefore necessary to be able to control the exposure and processing of the emulsion. Photographic film consists of the emulsion itself, and also the supporting base, and it is their combined amplitude and phase effect which produces the distribution of light in the diffraction plane.

II.1 Characteristic Curve

The photographic system can be considered as a two stage process. The first stage is the exposure producing the latent image, the second stage is the development of the latent image to produce the final

image. It is a well known fact that the photographic process is only linear over a limited range, it is therefore necessary first to establish the conditions of exposure and development for which the properties of linearity exist. This is essential because the emulsion will be used in systems where Fourier transform techniques are utilised and such techniques are only valid for linear systems. Hopkins[37] said that for a system to be linear it must satisfy two postulates; the first is that of superposition, i.e. the additivity of responses, the second is that of stationarity, i.e. the image forming properties are not position dependent. The second postulate is generally met in most photographic systems, especially if only small areas of the emulsion are used. The postulate of superposition is not however, valid for photographic systems mainly due to the adjacency effect and to the non-linear relationship between the exposure and response. When the photographic system is treated as a two stage process the first stage is simply optical diffusion and can be regarded as linear. The output from the second stage is a distribution of density or amplitude transmittance which can be measured. The characteristic curve for the emulsion is the relationship between the input, which is the exposure, and the output. Linear recording is achieved when we have a linear relationship between the input and the output. Measurements of the output can be described by three different readings, namely density, intensity transmittance and amplitude transmittance. There is of course a relationship between the three terms, it is not however a linear relationship and hence the condition for linear recording will depend on the method by which we wish to measure the output. It can easily be seen that the shape of the characteristic curve will be dependent on the development conditions

in the second stage of the process.

There are two means whereby an image, which is a distribution of light, can be recorded photographically :

(a) The classical method of forming an image in the plane of the emulsion by means of a lens, the information is then localised at a relevant point.

(b) The holographic method of recording the interference pattern produced by light coming from an object and a reference beam. The object can then be seen by reconstructing with the same reference beam. In this case the information is reproduced at every point on the whole surface of the hologram.

In conventional classical photography one usually works in incoherent illumination, whereas in holography one utilises coherent light sources. It is therefore reasonable to expect the behaviour of the photographic emulsion to be different in both methods.

In conventional photography the characteristic curve describes the relationship between optical density produced by a given exposure, where density expresses the amount of light absorbed. We use this relationship because it is meaningful when one considers that the eye is used to scan the photograph. Linearity for such a system is the straight line portion of the density-log exposure (D-log E) curve. The gradient of this part is known as the gamma (γ). In holographic coherent optical systems, we are not so much interested in the absorption as the attenuation of the light. Hence we speak in terms of amplitude transmittance rather than density. The characteristic curve for such a system is the amplitude transmittance-exposure relationship (Ta-E), because the holographic process requires linearity between the exposure and the resulting transmittance.[25,26,46,53,65]

We therefore require $T_a = kE$ where k is a constant now density is defined as $D = \log (1/T)$, T being the intensity transmittance.

Therefore we have $D = -2 \log T_a$ and $\gamma = \frac{dD}{d(\log E)}$

therefore $\gamma = \frac{d(-2\log T_a)}{d(\log E)} = \frac{-2E\,dT_a}{T_a\,dE}$

therefore $T_a = \frac{2E}{\gamma}\frac{dT_a}{dE} = kE$

where $k = -\frac{2}{\gamma}\frac{dT_a}{dE}$

for $\gamma = -2$ i.e. for a reversal k is the slope of the T_a - E curve. Hence linearity can be satisfied.

There is a considerable region for which approximate linearity is obtained on the T_a-E curve. Unfortunately this does not correspond to the linear region of the D-log E curve, hence optimum recording for conventional photography does not give optimum conditions for holographic recording.

It is sometimes more convenient to speak in terms of log E rather than E[5,40]. Biedermann shows that the gradient of the amplitude transmittance-log exposure curve is important with regard to luminance of reconstruction in coherent optical systems. The maximum luminance does not coincide with the linear region of the T_a-E curve.

Examples of three possible types of characteristic or transfer curves are shown in Fig. 1. The shape of the characteristic curve for a particular set of exposures will depend on the developer, developing time and the temperature of development. Hence it is first necessary to vary these conditions to find the best curve for linear response.

II.2 Holographic Recording and Linearity

Let us consider the general case of two coherent spherical waves

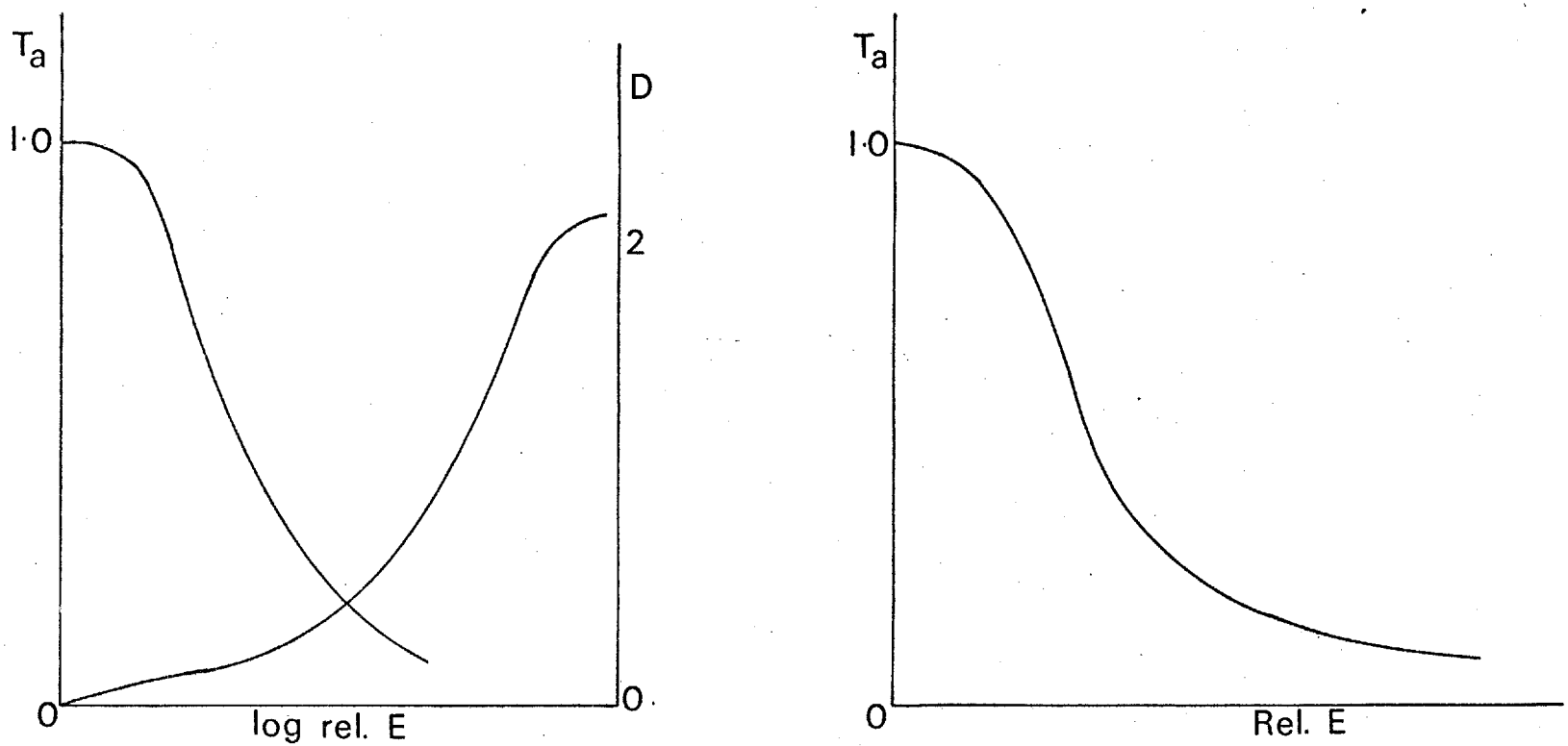


Fig 1 Characteristic Curves

arriving at the hologram surface H, (Fig. 2).

The amplitude at N due to a point b on the object is of the form $b \exp(i2\pi x \sin\phi_2 / \lambda)$ and the amplitude from the reference beam is a $\exp(i2\pi x \sin\phi_1 / \lambda)$ a, b being the respective amplitudes, ϕ_1, ϕ_2 the angles formed from the normal to the hologram at N, and λ the wavelength of light used. For a small pencil of rays we can assume the spherical waves are plane waves.

The resultant intensity in the plane of the hologram is then :

$$I = a^2 + b^2 + 2ab\cos\left[\frac{2\pi x}{\lambda}(\sin\phi_1 - \sin\phi_2)\right] \qquad 2.1$$

Therefore we now have a sinusoidal variation of intensity on the hologram surface where the average or bias level is $a^2 + b^2$ and the frequency of the intensity variations is given by :

$$u = (\sin\phi_1 - \sin\phi_2)/\lambda \qquad 2.2$$

If A and B are the intensities due to the reference beam and object point at the hologram surface, and θ the angle between the beams, then the intensity becomes :

$$I = A + B + 2(AB)^{1/2}\cos\left\{\frac{2\pi x}{\lambda}[\sin(-\phi_1) - \sin(\theta - \phi_1)]\right\}$$

$$= A + B + 2(AB)^{1/2}\cos\left[\frac{4\pi x}{\lambda}\sin\frac{\theta}{2}\cos\left(\phi_1 - \frac{\theta}{2}\right)\right] \qquad 2.3$$

For normal symmetrical recording conditions $\theta = 2\phi_1$, then the frequency of the fringes is given by :

$$u = \frac{2}{\lambda}\sin\frac{\theta}{2} \qquad 2.4$$

If however, the hologram is tilted with respect to the two beams and $\theta \neq 2\phi_1$, then the fringes are modified by a factor $\cos(\phi_1 - \frac{\theta}{2})$. If

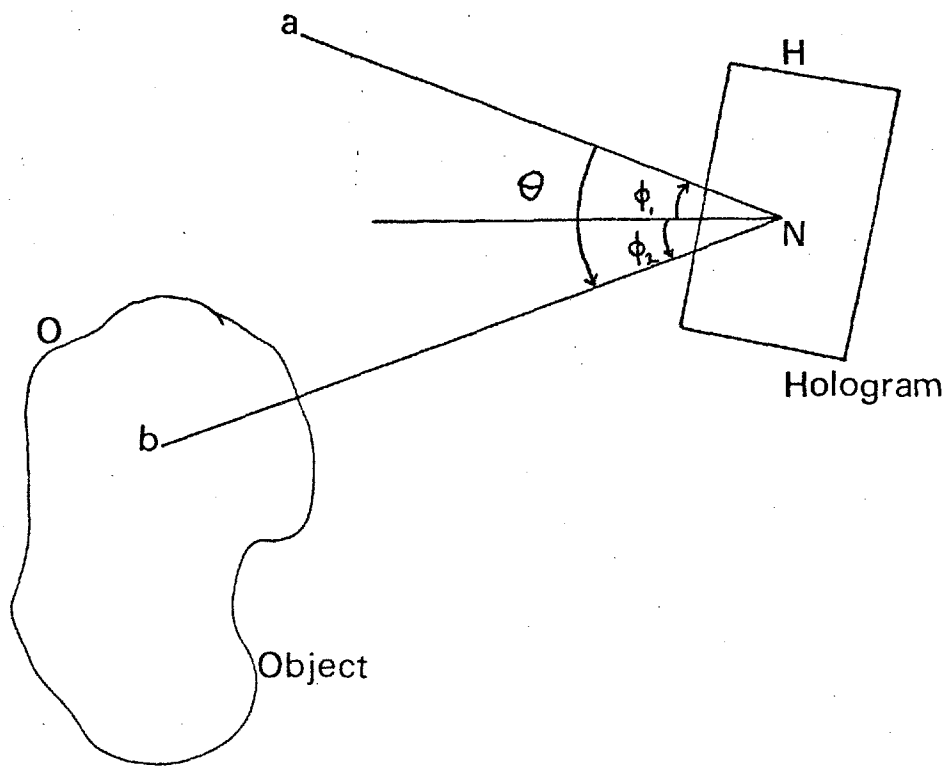


Fig 2 Hologram Recording

we apply 2.3 to the formation of a simple grating produced by two parallel beams I_o, I_s, intersecting normally at the hologram surface; then the intensity distribution of the interference pattern so formed is :

$$I = I_o + I_s + 2(I_o I_s)^{1/2} \cos\left(\frac{2\pi}{u} x\right) \qquad 2.5$$

u being the period of the fringes in the x direction.

In order to record this modulation :

$$u = \frac{2}{\lambda} \sin\frac{\theta}{2} < u_1$$

u_1 being the highest spatial frequency the emulsion can record for the specific developing conditions, and :

$$I_{max} = I_o + I_s + 2(I_o I_s)^{1/2} < E_2$$

$$E_1 < I_o + I_s - 2(I_o I_s)^{1/2} = I_{min}$$

E_1 being the threshold or minimum exposure for linearity and E_2 the saturation or maximum exposure for linearity.

The value of the exposure time is chosen such that the average intensity $I_o + I_s$ occurs at the centre of the region $E_1 - E_2$ of the T_a-E curve, this being the bias point.

The modulation or contrast of the fringes is given by :

$$M = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{2(I_o I_s)^{1/2}}{I_o + I_s} = \frac{2(I_o/I_s)^{1/2}}{1 + I_o/I_s} \qquad 2.6$$

For linear recording we choose the ratio $I_o : I_s$ Equation 2.6 such that the modulation about the bias point remains in the linear region. The amplitude transmittance of this grating after exposure and development is :

$$T_a = T_b + 2k(I_o I_s)^{1/2} \left[t K\left(\frac{1}{u}\right)\right] \cos\left[\frac{2\pi x}{u} - \phi_1\left(\frac{1}{u}\right)\right] \qquad 2.7$$

where T_b is the amplitude transmittance at the bias point due to the

average intensity $I_o + I_s$, k is the slope of the amplitude transmittance-exposure curve at the bias point, and K and ϕ are dependent on the spatial frequency recorded and define the amplitude and phase of the optical transfer function of the emulsion. Linear amplitude recording is achieved when a sinusoidal varying input signal onto the emulsion produces after development an undistorted sinusoidal amplitude transmittance variation.

It is therefore seen that linear recording can be achieved by two means;

(a) by having a small modulation about a bias point or in fact having phase modulation only where the phase modulation is linear with exposure. This however would result in low contrast recording which does not give the best reconstruction. Many systems optimise for the best contrast with minimum non-linear effects.

(b) by processing to a positive hologram transparency having the proper gamma characteristics. This, suggested by Gabor[30] would require making the gamma of the two processes equal to 2 and would place severe restrictions on the exposure and development.

To show this let T_a = amplitude transmittance as before and T_i to be the intensity transmittance. The exposure $E(x,y) = E_b + E_s(x,y)$ where E_b is the bias and E_s is an A.C. varying signal. Then :

$$T_i(x,y) = \left[\frac{E(x,y)}{E_o}\right]^{-\gamma} \qquad 2.8$$

E_o being the intercept of the tangent at the bias point on the D-log E curve. Then :

$$T_i(x,y) = \left[\frac{E_b + E_s(x,y)}{E_o}\right]^{-\gamma} \qquad 2.9$$

$$= \left(\frac{E_b}{E_o}\right)^{-\gamma}\left[1 + \frac{E_s(x,y)}{E_b}\right]^{-\gamma} \qquad 2.10$$

now exposure = intensity x time :

$$\text{therefore} \quad T_i(x,y) = T_{iB}\left[1 + \frac{I_s(x,y)}{I_b}\right]^{-\gamma} \qquad 2.11$$

where T_{iB} is the bias of the intensity transmittance.

For linearity we require $T_a \propto E$ which is the same at $T_a \propto I_s$

now $T_i = |T_a(x,y)\, e^{i\phi(x,y)}|^2$

where $e^{i\phi(x,y)}$ is the phase effect of the emulsion due to thickness changes which are proportional to density D, See 2.7 and Falconer[22].

$$\text{therefore} \quad T_a = T_i^{1/2}\, e^{-i\phi(x,y)}$$

$$= T_i^{1/2}\, e^{-iKD} = T_i^{1/2 + iK'} \qquad 2.12$$

let m = transfer function of the emulsion. Then from 2.11 and 2.12

$$T_a = \left[T_{iB}\left(1 + m * \frac{I_s(x,y)}{I_b}\right)^{-\gamma}\right]^{(1/2 + iK')}$$

where * denotes convolution

$$\text{therefore} \quad T_a = T_{iB}^{1/2 + iK'}\left[1 + m * \frac{I_s(x,y)}{I_b}\right]^{-\gamma(1/2 + iK')} \qquad 2.13$$

Making a positive transparency from this negative would give an amplitude transmittance of :

$$\bar{T}_a = \bar{T}_{iB}^{1/2 + iK'}\left[1 + m * \frac{I_s}{I_b}\right]^{\gamma\bar{\gamma}(1/2 + iK')} \qquad 2.14$$

for linearity $T_a \propto I_s$. Hence from 2.14 $\gamma\bar{\gamma}\,\frac{1}{2} = 1$, which gives the result that the gamma of the two processes should be equal to 2 for linear recording.

II.3 Non-Linearity

Although it is possible to record linearly using a photographic emulsion it is not always the most efficient way of using the emulsion. The response of a non-linear photographic process has been treated by

several authors[14,26,32,47,48,49,62,71,77,86], and is mentioned in numerous other papers. All of these studies result in obtaining some kind of general expression for the higher order images produced by non-linear recording. They start from the basic T_a-E characteristics and by various means try to obtain some form of polynomial to the non-linear terms. The general overall result is to produce undesired orders in the reconstruction process together with shape distortion and image blurring resulting in a loss of diffraction efficiency.

II.4 Dynamic Range and Noise

The dynamic range can be defined as the ratio of the intensity of light produced by the largest amplitude two dimensional sinusoidal signal which can be recorded on the emulsion, to the intensity of light due to noise which falls within a unit bandwidth at the spatial frequency of the recorded sinusoidal signal.

The dynamic range is therefore the number of intensity levels per unit bandwidth that can be distinguished by the developed emulsion. The noise arises due to the grainy structure of the emulsion with the result that when the emulsion is developed to a constant macroscopic transmittance, there exists a microscopic complex transmittance which varies from point to point on the emulsion. Light incident onto this emulsion is diffracted or scattered by it and interferes with the signal recorded on the emulsion. The result is that the noise, which is wide band, limits the dynamic range of the emulsion. Noise generated by the granular structure of the emulsion is entirely random in nature. The subject has been treated by many authors[1,6,20,33.34,35] and will not be dealt with here. Scattering by the granular structure will be dealt with in Chapter V.

To evaluate the dynamic range of a particular emulsion a

sinusoidal varying signal is exposed onto the emulsion. If the signal is of the form of cos $(2\pi ax+2\pi by)$ then the amplitude transmittance is given by 2.13.

$$T_a(x,y) = T_{ib}\left[1 + m(x,y) * \cos(2\pi ax + 2\pi by)\right] \qquad 2.15$$

for linear recording of maximum amplitude sinusoidal signal and no phase effects present. If this transparency is then placed in the signal plane of a spectrum analyser as described in the previous chapter, then the intensity in the frequency plane is given by the modulus of the square of the Fourier transform of the amplitude transmittance of the transparency.

$$I(\xi,\eta) = \left| \iint_{-L}^{L} T_{ib}^{1/2}\left[1 + m(x,y) * \cos(2\pi ax + 2\pi by)\right] e^{-i2\pi(a_o x + b_o y)} \cdot dx\,dy \right|^2$$

$$= T_{ib}\left(L^2 \operatorname{sinc}\pi L a_o \operatorname{sinc}\pi L b_o\right)^2 \qquad 2.16$$

$$+ T_{ib}\left[\tfrac{1}{2}\bar{m}(a_o b_o) L^2 \operatorname{sinc}\pi L(a_o - a) \operatorname{sinc}\pi L(b_o - b)\right]^2$$

$$+ T_{ib}\left[\tfrac{1}{2}\bar{m}(a_o b_o) L^2 \operatorname{sinc}\pi L(a_o + a) \operatorname{sinc}\pi L(b_o + b)\right]^2$$

The first term of 2.16 is the bias intensity transform, the second term the positive frequency image and the third term the negative frequency image. The total intensity in a unit bandwidth about the positive frequency is given by :

$$T_{ib}\left[\tfrac{1}{2}\bar{m}(a_o b_o) L^2 \tfrac{1}{2}\right]^2 . L^{-2} = T_{ib}\left[L\bar{m}(a_o b_o)\tfrac{1}{4}\right]^2 \qquad 2.17$$

To find the corresponding noise intensity a piece of similar emulsion is uniformly exposed and developed to a mean intensity transmittance T_{ib} of the previous sample. This is now placed in the signal plane of the spectrum analyser previously occupied by the sinusoidal signal. The macroscopic amplitude transmittance of this sample is :

$$T_a(x,y) = \left[T_{ib} + T_{in}(x,y)\right]^{1/2} \approx T_{ib}^{1/2}\left[1 + \tfrac{1}{2}\frac{T_{in}(x,y)}{T_{ib}}\right] \qquad 2.18$$

Hence the noise intensity falling in a unit bandwidth about the positive frequency is :

$$\frac{1}{4} T_{ib} \left[\frac{T_{in}(a_c b_c)}{T_{ib}} \right]^2 L^{-2} \qquad 2.20$$

The dynamic range of the emulsion is then found from 2.17 and 2.20.

$$\text{Dynamic range} = \left[\tfrac{1}{2} L^2 \bar{m} (a_c b_c) \right]^2 \Big/ \left[\frac{T_{in}(a_c b_c)}{T_{ib}} \right]^2 \qquad 2.21$$

Hence it is seen that the dynamic range decreases with increasing noise. The ratio of T_{in}:T_{ib} can be expressed in terms of the granularity of the emulsion[22].

II.5 Diffraction Efficiency

Coherent light incident onto a grating is diffracted, the diffraction ocurring being a function of the grating. A hologram is a form of grating, it is formed by the interference of two beams producing spatially varying amplitude and phase distributions. It would therefore be reasonable to assess the performance of a hologram by considering the amount of diffracted light which contributes to the reconstruction of the original wavefront. The diffraction efficiency is defined as the ratio of light intensity diffracted by the hologram producing the original wavefront in reconstruction to the intensity of light incident on the thin hologram.

The interference pattern recorded can take various forms depending on the medium used. So far all the analysis has been performed assuming the interference produced results in a two dimensional amplitude transmittance variation. This is known as the thin transmittance grating. If this film is bleached the hologram is recorded by spatially varying phase information and is a thin phase

grating. If however, the thickness of the emulsion is many times greater than the wavelength of light used then it can no longer be regarded as a two dimensional storage medium, but as a three dimensional, (thick) volume hologram. As before this can either be an amplitude or a phase grating depending upon the way the diffraction occurs. Holograms formed in such a thick medium can be either of the transmitting or reflecting type, depending on whether the reconstruction is viewed in transmission or reflection, the difference between the two being that in a transmission hologram the orientation of the fringes is essentially perpendicular to the surface of the emulsion, whereas in a reflection hologram the orientation is essentially parallel to the emulsion surface. Furthermore under many circumstances practical holograms affect both the amplitude and phase simultaneously hence they are strictly speaking complex holograms. Kogelnik[43] compares the response of these various hologram types in their reconstruction process. Similarly other authors[4,12,15,27,28,31,44,54,58,66,69,82,83,84] discuss the reconstructing capabilities of the various holograms. We will restrict ourselves to the performance of the thin hologram.

For these types of holograms the two dimensional light intensity of the interference pattern on the emulsion surface is given by Equation 2.5. Assuming this signal is linearly recorded then Friesem, Kozma and Adams[25] show that the intensity of the diffracted light due to the amplitude grating is :

$$I_d = (\beta t m)^2 \, I_o I_s I_g$$

where β is slope of the T_a-E curve, t the exposure time, m the modulation transfer function of the film and I_g the intensity of light transmitted by an unexposed portion of the film. The normalised diffracted light

intensity or diffraction efficiency is :

$$I_d/I_g = (\beta t m)^2 \; I_o I_s \qquad 2.22$$

The square root of $I_o I_s$ is linearly related to the modulation 2.6, hence the diffraction is linearly related to the square of the modulation at a fixed bias point. In their paper, Friesem et al investigated the effect of various recording systems on the diffraction efficiency. They found :

(a) the diffraction efficiency was maximised for an amplitude transmittance of 0.5 at the bias point, this coincided with the centre of the linear portion of their T_a-E curve.

(b) the square root of the diffraction efficiency increased linearly with the modulation at a fixed bias point and spatial frequency.

(c) the diffraction efficiency decreased with increasing spatial frequency of fringes recorded at a fixed bias point and modulation.

It can be seen from Equation 2.17 that the maximum diffraction efficiency Idiff/Iinc of such an amplitude grating is 1/16 or 6.25%.

A thin phase hologram can be obtained either by bleaching a thin amplitude grating or recording on such a medium that the exposure produces directly a change of optical thickness. The signal is now recorded as a spatially varying phase shift. The amplitude transmittance is given by :

$$T_a(x) = 1 . e^{-i\phi(x)}$$

where $\phi(x) = \phi_c - \phi_1 \cos 2\pi u x$, u being the frequency of variations.

This type of grating will diffract an oncoming plane wave into a series of equally spaced spectra. The amplitudes of the spectra are obtained from the Fourier transform of the spatially varying transmittance

term $\phi_1 \cos 2\pi ux$, and are :

$$\sum_{-\infty}^{\infty} i^n J_n(\phi_1)\, e^{in2\pi ux}$$

where J_n is a Bessel function of order n.

The diffraction efficiency is found from either of the first order spectra which contribute to the reconstruction of the original wavefront. The second and higher order spectra add unwanted distortion to the reconstructed beam. The diffraction efficiency of such a grating is given by $J^2_{1max} = 33.9\%$.

Linear recording is achieved when the exposure is linearly related to the phase variations produced.

II.6 Optical Transfer Function (O.T.F.)

The photographic emulsion is a medium used to store an image. It is therefore necessary to know the imaging properties of the emulsion. The optical transfer function gives us a relationship between the contrast of the object and the recorded image. It can be defined as the ratio between the image contrast and the object contrast or the ratio of the Fourier spectrum of the image to that of the object. The o.t.f. can be measured using a periodic object such as a sine wave grating or by a non-periodic object as a slit or an edge. It basically tells us how a system can record various frequencies and is therefore sometimes called the frequency response.

We have already stated that application of Fourier techniques requires a linear system. The measurement of o.t.f. requires the use of Fourier techniques, hence it is only meaningful to speak of the o.t.f. of an emulsion if the recording is linear or if the results are corrected for the non-linearities of the photographic process. The two stage process has a linear first stage producing the latent

image, therefore a sinusoidal exposure gives a sinusoidal latent image. The second stage, however, is non-linear and the transmittance is not sinusoidal. It can be expressed by a fundamental frequency, equal to the incident exposure frequency, together with a series of higher harmonics. The transfer characteristics of the photographic system such as the film and subsidiary optics, have characteristics of low pass filters, thus reducing the amplitude of the harmonics and hence the output signal tends to be sinusoidal.

Let us first consider the image of a line on a photographic emulsion under linear conditions. The input intensity to the emulsion is given by B(x) and the output by B'(x'). The relationship between the two quantities is given by :

$$B'(x') = \int_{-\infty}^{\infty} B(x)\, G(x'-x)\, dx \qquad 2.23$$

$G(x'-x)$ being the line spread function. The image, object and line spread function are related by a convolution integral. Suppose the exposure is a sinusoidal varying signal $B(x) = a+b.\cos(2\pi ux)$, u being the frequency of the fringes in the x direction.

Then

$$B'(x') = \int_{-\infty}^{\infty} [a+b\cos(2\pi ux)]\, G(x'-x)\, dx \qquad 2.24$$

Let $x'-x = p$ then the intensity at x' is :

$$B'(x') = \int_{-\infty}^{\infty} [a+b\cos\{2\pi u(x'-p)\}]\, G(p)\, dp \qquad 2.25$$

Let $C(u) = \int_{-\infty}^{\infty} G(p)\cos(2\pi up)\, dp$

and $S(u) = \int_{-\infty}^{\infty} G(p)\sin(2\pi up)\, dp$

Then

$$B'(x') = C(0)\left\{a+b\left[\frac{C(u)^2+S(u)^2}{C(0)}\right]^{1/2}\cos[2\pi ux'+\theta(u)]\right\} \qquad 2.26$$

where $\theta(u) = \arg[C(u)-S(u)]$, $C(0)$ being a constant. The Fourier transform

of of the line spread function is :

$$g(u) = \int_{-\infty}^{\infty} G(x)\, e^{i 2\pi u x}\, dx$$

$$g(u) = C(u) - i\,S(u)$$

The modulation or contrast transfer function $T(u) = \left|\frac{g(u)}{g(o)}\right|$

Then from 2.26 :

$$B'(x') = a + b\,T(u)\cos[2\pi u x' + \theta(u)] \qquad 2.27$$

The image of a sine wave grating is therefore a sine wave grating with reduced contrast and phase shift.

Now input contrast = $\frac{b}{a}$, image contrast = $\frac{bT(u)}{a}$ therefore T(s) = $\frac{\text{image contrast}}{\text{object contrast}}$.

The transfer function of the emulsion can then be found by scanning a series of known modulation sinusoidal signals with a microdensitometer and evaluating the modulation reocrded. It could also be measured by scanning the line spread function with a slit and obtaining the Fourier transform of the intensity distribution. The theory described above refers to imaging of an incoherent object, since we have been using intensities. For coherent image formation we must consider amplitude instead of intensity.

For coherent illumination let us define the transfer function by K(x,y). Now the relationship between the image, object and transfer function is as before given by a convolution integral :

$$u'(x', y') = \int_{-\infty}^{\infty} u(x,y)\, K(x'-x, y'-y)\, dx\, dy$$

and by using the convolution theorem we obtain

$$u'(u,v) = u(u,v)k(u,v)$$

where (u,v) represent spatial frequencies. This shows us that the spatial frequency content of the complex amplitude is transferred from the object to the image according to the action of a linear filter.

It can also be shown[8] that for coherent illumination the transfer function k(u,v) is equal to the value of the pupil function G at that point of the Gaussian reference sphere. Since the value of an unobstructed pupil function is constant within the boundary of the aperture and zero outside then the value of the transfer function for a perfect system is unity up to the limit of the frequencies transmitted. For a circular aperture of radius a the frequency cut off is given by :

$$u^2 + v^2 = \left(\frac{a}{\lambda R}\right)^2 \quad .$$

R being the radius of the reference sphere.

For incoherent illumination the transfer function is given by the auto-correlation function of the pupil function. As we displace the apertures relative to each other the value of the transfer function is given by the area common to both apertures. Hence, for the incoherent case the value of the transfer function decreases with increasing spatial frequency and has a cut off when the two areas just do not overlap. For a circular aperture this happens when the centres of the apertures are separated by a distance of 2a. The cut off frequency is then given by :

$$u^2 + v^2 = \left(\frac{2a}{\lambda R}\right)^2 \quad .$$

Hence it is seen that the spatial frequency cut off for incoherent illumination is twice that for coherent illumination for the same optical configuration. Fig. 3 shows the variation of the transfer function for the two cases for a perfect system.

Friesem, Kozma and Adams related the modulation transfer function to the diffraction efficiency. Using Equation 2.22 :

$$D\left(\frac{I_d}{I_g}\right)^{1/2} = \frac{2(I_o I_s)^{1/2}}{I_o + I_s} \qquad (2.28)$$

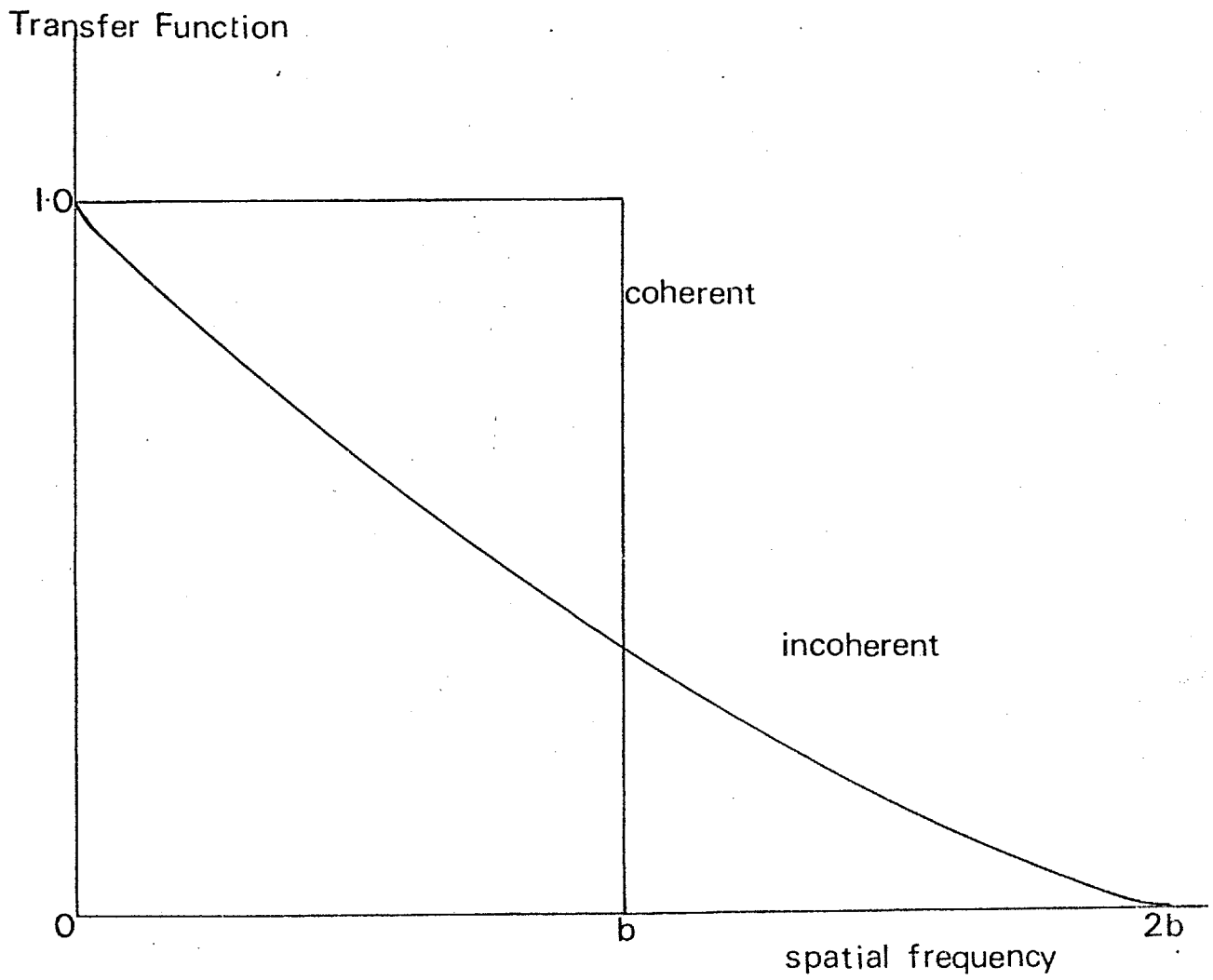


Fig 3 Optical Transfer Function

where D is a proportionally constant. From measurements of light diffracted into one of the first orders of a sinusoidal grating :

$$\left(\frac{I_d}{I_g}\right)^{1/2} = (\beta t m)(I_o I_s)^{1/2} = \frac{\Delta T}{2} \qquad (2.29)$$

where ΔT is the ratio of the grating's average to peak amplitude transmittance variation. Determination of the same ratio for a low frequency grating can be found similarly by making a low frequency grating, so that in this case :

$$\left[\frac{I_d}{I_g}\right]^{1/2}_{LF} = \left[\beta t (I_o I_s)^{1/2}\right]_{LF} \qquad (2.30)$$

The transfer function m is : $\dfrac{(I_d/I_g)^{1/2}}{(I_d/I_g)^{1/2}_{LF}}$

which is the diffracted amplitude normalised with respect to a relative low frequency diffracted amplitude.

The main factor influencing the o.t.f. of a photographic emulsion apart from its own natural limitation, is the adjacency effect. This occurs when the chemical changes produced in the developer by the reduction of the silver halide grains change the local development rate in the emulsion layer. The effect could be to either retard or accelerate the development at that part of the emulsion with the result that the density produced would be lower or higher than expected for that exposure. This in turn would affect the contrast and hence the o.t.f. of the emulsion. Adjacency effects can be minimised by continuous agitation during development to distribute the reaction products uniformly, and by using large areas to minimise the small local area effects.

II.7 Phase Effects

When using photographic film in a coherent optical system we are

interested in its complex amplitude transmittance. If it is to be used as an amplitude grating then the phase effects could tend to distort the frequency spectrum. The phase effects can arise from two sources. Firstly, film is not optically flat even when exposed to a uniform density. Secondly, phase effects arise due to the fact that the thickness of the developed film varies with its density[2,50,72,73]. Both of these effects are undesirable for film used in data processing, for in such cases the output is of the Fourier transform of the complex amplitude transmittance of the film and would be distorted. The first type of phase effects are random in nature and will be dealt with experimentally in a later chapter. The second type is dealt with below.

During the development stage of the photographic process the reaction products of some developers tan the gelatin, these parts then absorb less water than the untanned portions. In an aqueous solution therefore the gelatin swells, the extent of which depends on the degree of tanning. The degree of tanning is proportional to the reaction products produce which in turn is proportional to the quantity of silver reduced. Hence the amount of swelling is proportional to the image density. During the drying process the more heavily tanned (high density) areas dry more quickly than the less tanned areas and pull these low density areas towards them. The resulting effect is a change of thickness as well as a possible phase shift in the emulsion. It is the changes of thickness which produce the variations in phase. They are called relief heights. When referring to density and transmittance we are talking about their diffuse and not specular values.

The phase retardation produced by the relief heights is a linear function of the density (this assumption will be verified experimentally

in Chapter V), therefore phase retardation = kD where k is a constant of the relief height and D the density. Let us consider the effects of these relief heights on a sinusoidally exposed grating. The complex amplitude transmittance T_c is given by :

$$T_c = T_a e^{i \frac{2\pi}{\lambda} kD} \tag{2.31}$$

$T_a = a + b \sin\theta$ is the amplitude transmittance. Now $D = -\log_{10} T$

therefore : $T_c = T_a e^{-ik\frac{2\pi}{\lambda} 2 \log_{10}(a + b\sin\theta)}$

$$= T_a \exp\left[-ik\frac{4\pi}{\lambda} 0.434 \log_e a\left(1 + \frac{b}{a}\sin\theta\right)\right] \tag{2.32}$$

$$= T_a \exp\left\{-ik\frac{4\pi}{\lambda} 0.434\left[\log_e a + \log_e\left(1 + \frac{b}{a}\sin\theta\right)\right]\right\} \tag{2.33}$$

Let $z = \frac{4\pi}{\lambda}$ k 0.434 $\log_e a$ then 2.33 becomes :

$$T_c = e^{-iz} T_a e^{-iz \log_e(1 + b/a \sin\theta)}$$

$$= e^{-iz} T_a \left\{\cos\left[z \log_e\left(1 + \tfrac{b}{a}\sin\theta\right) - i\sin\left[z \log_e\left(1 + \tfrac{b}{a}\sin\theta\right)\right]\right\}\right. \tag{2.34}$$

for $b/a \ll 1$ $\log_e(1 + b/a \sin\theta) \approx b/a \sin\theta$

Using this in 2.34 : $T_c = e^{iz} T_a \left[\cos\left(z \frac{b}{a}\sin\theta\right) - i\sin\left(z \frac{b}{a}\sin\theta\right)\right]$ (2.35)

now $\cos(x\sin\theta) = J_0(x) + 2J_2(x)\cos 2\theta + 2J_4(x)\cos 4\theta + \ldots 2J_n(x)\cos n\theta$

and $\sin(x\sin\theta) = 2J_1(x)\sin\theta + 2J_3(x)\sin 3\theta + \ldots 2J_n(x)\sin n\theta$

Jn being the nth order Bessel function, using this for 2.35 and with

$x = z\frac{b}{a}$

therefore $T_c = e^{-iz(a + b\sin\theta)}\left[J_0(x) + 2J_1(x)\cos 2\theta + \ldots - i\left(2J_1(x)\sin\theta + 2J_3(x)\sin 3\theta + \ldots\right)\right]$

This gives the amplitude of the zero order term.

$$= a J_0\left(k \frac{0.434}{\lambda} 4\pi \frac{b}{a}\right) - i b J_1\left(k \frac{0.434}{\lambda} 4\pi \frac{b}{a}\right) \tag{2.36}$$

and the total amplitude of the first order spectra

$$= b\left[J_0\left(k \tfrac{0.434}{\lambda} 4\pi \tfrac{b}{a}\right) - J_2\left(k \tfrac{0.434}{\lambda} 4\pi \tfrac{b}{a}\right)\right] - i 2a J_1\left(k \tfrac{0.434}{\lambda} 4\pi \tfrac{b}{a}\right) \tag{2.37}$$

These functions of 2.36 and 2.37 can be plotted in the complex plane. In theory the value of x can go from zero to infinity, this however does

not occur in practice. Figs. 4 and 5 show the zero and first order spectra plotted in the complex plane for k increasing and $b = a = \frac{1}{2}$. The amplitude of the spectra is given by the length of a line from the origin to the point in question on the curve. From Fig. 4 it is seen that the zero order is a maximum when there is no phase modulation and it becomes smaller as the phase modulation increases. The vector of the amplitude generates a spiral. Fig. 5 is similar in shape to Fig. 4. It is seen however that there appears to be a certain phase modulation which marginally increases the amplitude after which it decreases with phase modulation. It must be remembered that these curves have been plotted for $b/a = 1$ whereas the theory used to derive the curves assumes $b/a \ll 1$, hence the validity of the marginal increase of amplitude of the first order spectra could be questioned. For smaller modulation signals the zero order spectrum is compressed in the $J_1(x)$ direction and the first order spectrum in the $J_o(x) - J_2(x)$ direction. The result is that it is much more evident that is is possible to enhance the first order spectrum by having phase modulation. Figs. 6 and 7 show similar effects when $z = 4$ and the modulation goes from zero to one.

These types of phase effects cause some interesting results. Tiziani et al[78] in their work on the classification of grains of photographic emulsions by shape needed to know the Fourier transform of various shapes, one of these being a triangle. The Fourier transform of a triangle should show sixfold symmetry. In their work, however, they found that the triangle showed only threefold symmetry. They found that the relief image effect near the edge of the triangle would increase the optical thickness by about one half of a wavelength and it is due to this that they observed threefold symmetry for a triangle.

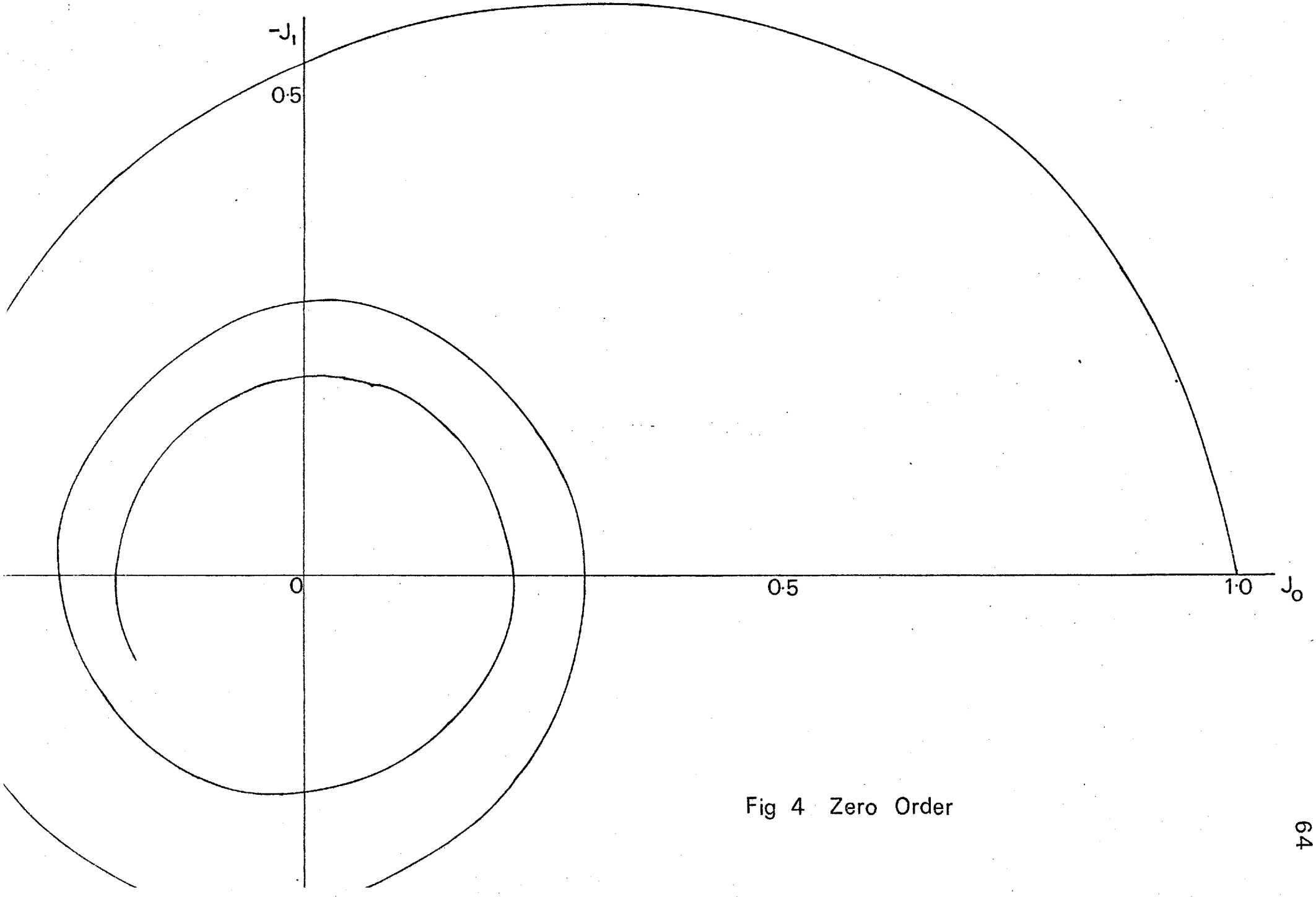


Fig 4 Zero Order

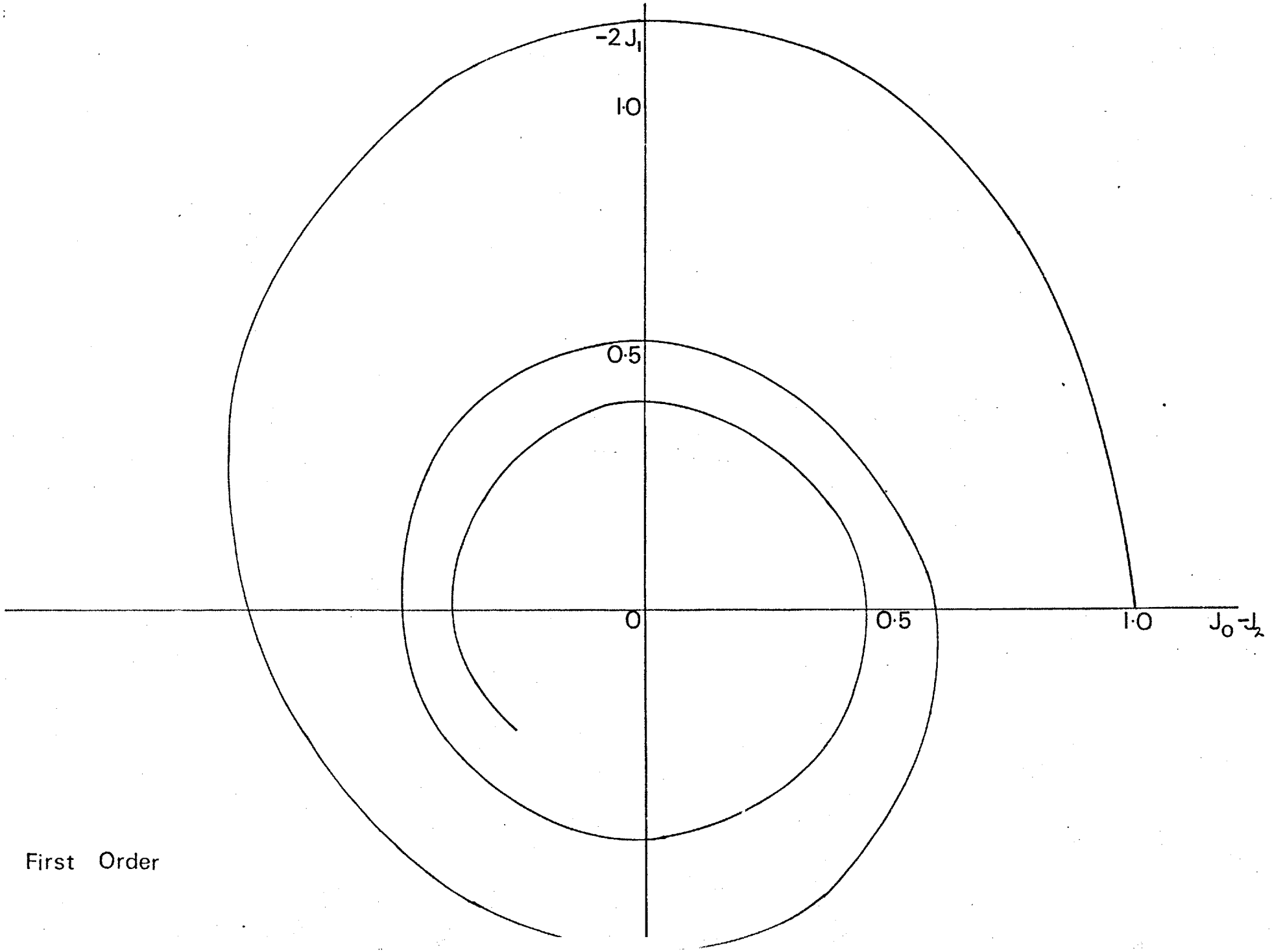


Fig 5 First Order

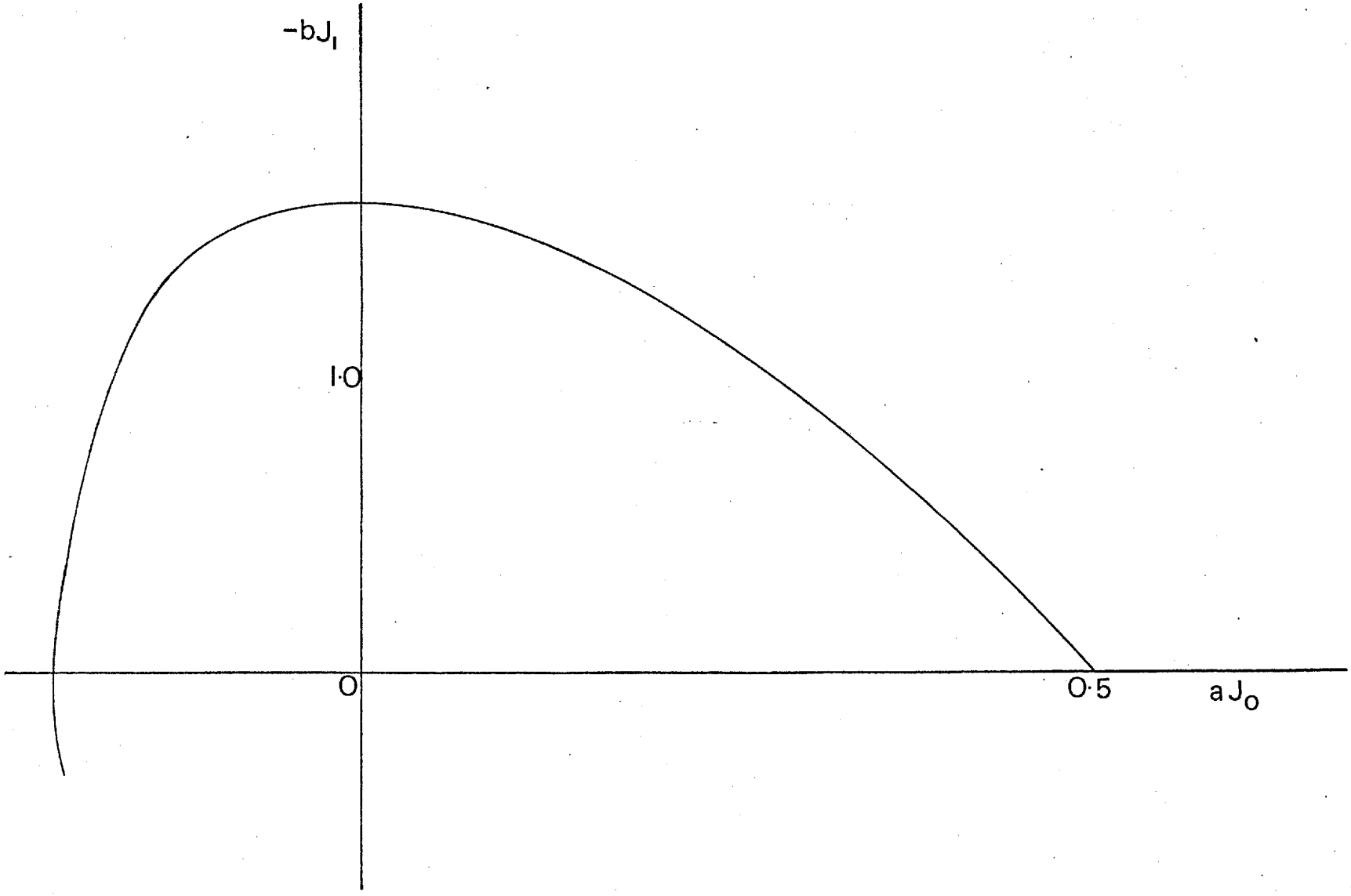


Fig 6 Zero Order $z = 4$

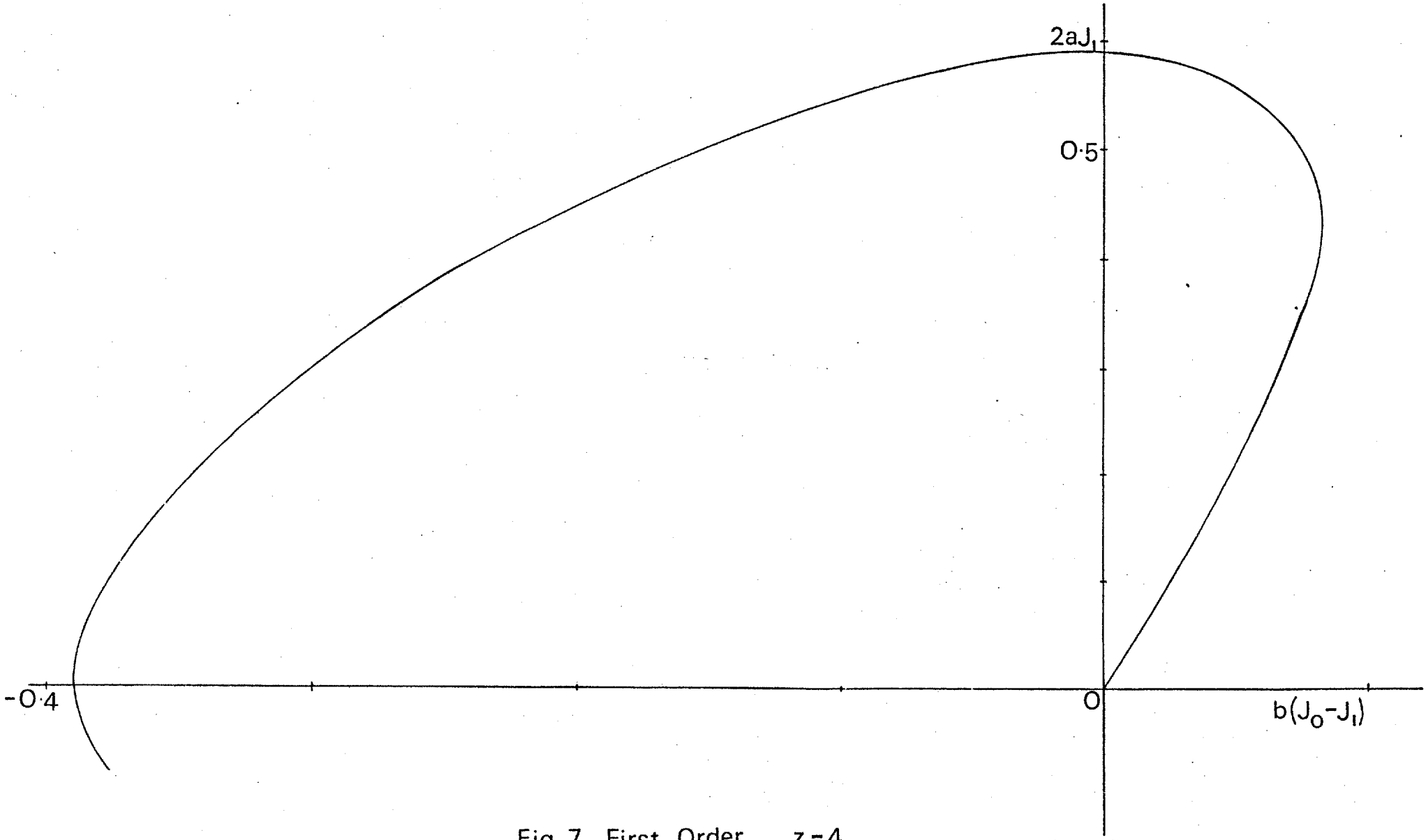


Fig 7 First Order z = 4

Thompson[76] attributes the degradation at the edges of a bar target noticed by Considine to the phase effects in the film formed during development.

These phase distortions can be eliminated by using a liquid gate. This consists in placing the film between two optical flats with a liquid of matching refractive index. Leith[52] demonstrated the use of such a liquid gate. He measured the relative intensities of the diffracted orders produced by a photographed sinusoidal signal. He then placed the film into a liquid gate and repeated the measurements. In all cases he found that the spectral content had been changed by removing the phase modulation. In general, energy was transferred from the higher diffracted spectra to the lower order ones, the zero order being increased up to 25 fold. For the case of a low modulation signal for which no higher orders occured (linear recording) he found that using a liquid gate decreased the first order spectra (and hence the diffraction efficiency). This agrees with the mathematical model that in certain cases the first order spectra are enchanced by phase modulation. In fact it has already been shown that using a pure phase grating the efficiency of the emulsion can be increased from 6.25% (thin amplitude grating) up to 33.9%.

CHAPTER III - THE MACH-ZEHNDER INTERFEROMETER

In the previous chapter we have indicated that the phase properties of photographic emulsions must be taken into consideration when the emulsions are being used with coherent radiation. Interferometers are useful instruments for examining phase variations produced in optical systems. For this purpose a Mach-Zehnder interferometer was constructed.

In any two-beam interferometer utilising an extended source the visibility or contrast of the fringes is highest at the position of localisation of the fringes, and it is also dependent on source size. As some of the emulsions being examined would have a density of over two a large source size would be advantageous if conventional light sources were used. Another important factor arising from the limited light intensity is that the beam should only pass through the sample once (unlike Twyman-Green interferometers). As regards fringe localisation the interferometer should be such that it would be possible to localise the fringes on the test sample so that the test sample and fringes could both be in focus. These two requirements are satisfied by the Mach-Zehnder interferometer.

III.1 Coherence

The radiation emitted from a light source can be considered from an elementary standpoint to be of the form of simple harmonic waves. The phase of the light at a point x depends upon the phase of the source and the distance of the point x from the source. If light beams from two different sources arrive at x there does not exist a fixed relationship between the phases of the two beams and they cannot

combine to form a stationary or interference wave. These light waves are called incoherent, the resultant intensity is the sum of the two individual intensites. If however, the radiation from one point of the source is split up and made to travel by two different paths to arrive at x a fixed phase relationship exists and the beams are called coherent. Superposition of such beams produce interference effects due to the combination of their amplitudes. Practical light sources however, do not consist of a single continuous wavetrain but of a series of pulses which are randomly related phasewise. This brings in a restriction to the difference of optical paths of the two beams forming the interference fringes.

Let us consider an atom of the source which emits a series of wavetrains. These wavetrains are made to fall upon a device which divides the incident wave into two waves which then follow two different paths before arriving at x. Therefore for each incident wavetrain we now have two wavetrains. If the difference of the two paths taken is greater than the length of the wavetrain then the wavetrains which meet at x came from two different wavetrains emitted by the source. As no correlation or phase relationship exist for different wavetrains no interference phenomena is observed. There is temporal incoherence due to the finite length of the wavetrains, which means that the emission from an atom of the source is not of a single frequency but rather a spectrum of frequencies. Temporal coherence is measured by the coherence time or coherence length of the source. The degree of temporal coherence depends upon the path difference between the two paths. If d is the path difference, and Δt the coherence time then for $d \ll c\Delta t$ no correlation exists between the two wavetrains and the mutual coherence is zero. The degree of mutual coherence increases as d

decreases and is maximum for d = o.

When the radiation has a range of frequencies the total intensity produced is proportional to the integral over all the frequencies. In a two beam interferometer the interference produced by a single frequency varies as $1 + \cos 2\pi \nu t$, ν being the frequency and t the time difference due to the path difference d of the two waves. If we have a range of frequencies, ν_o being the mean and $\Delta\nu$ the bandwidth then the intensity can be expressed as :

$$I(t) = G(0) + G(t)\cos 2\pi \nu_o t \qquad 3.1$$

The amplitude decreases with t thus limiting the time difference for which the fringes may be observed. The width of the function G(t) is related to the bandwidth by the uncertainty principle :

$$\Delta t \, \Delta\nu \approx 1$$

the coherence length is then $\Delta\ell = c\Delta t$, c being the velocity of light in the medium. The coherence length can thus be related to the bandwidth, expressed as a wave number by :

$$\Delta\lambda \, \Delta\sigma \approx 1 \qquad 3.2$$

From 3.2 we see that the temporal coherence increases as the bandwidth decreases. The ideal sources for an interferometer would therefore have a long coherence length which is associated with narrow natural line width. The temporal coherence is determined by the spectral characteristics of the source.

So far we have considered the wavetrains from a single atom. We must now consider an extended source with a narrow natural line width. Let S be the thermal source that illuminates two points P_1 and P_2, (Fig. 1). Due to random phase jumps the light from two separate atoms are incoherent. Consider atom a, which sends vibrations to P_1 and P_2. In general the distances a_1P_1 and a_1P_2 are not the same, hence

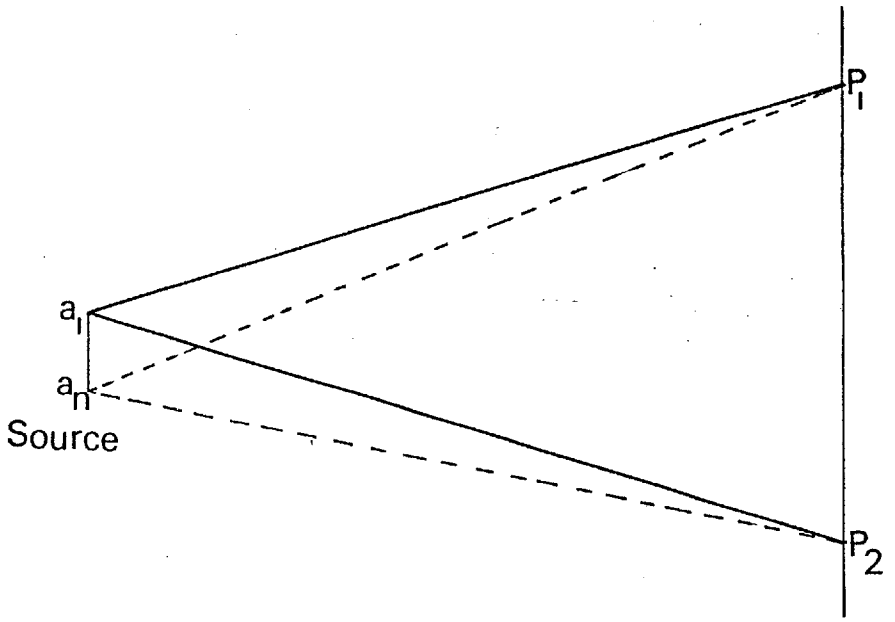


Fig 1 Extended Source

a phase difference of θ_1 occurs for the vibrations reaching P_1 and P_2 from atom a_1. Similarly for the other n atoms the vibration reaching P_1 and P_2 will have a phase difference of $\theta_2, \theta_3 \ldots \theta_n$ with no relationship existing between the θ's. If the path differences $a_1P_1 - a_1P_2$ are small compared to the wavelength then the phases $\theta_1 \ldots \theta_n$ are small and the points P_1, P_2 will receive vibrations from each atom which are practically in phase. For coherence we must have $[a_1P_1 - a_1P_2] - [a_2P_1 - a_2P_2] \ll \lambda$. The overall vibration states of P_1 and P_2 will be the same regardless of the number of atoms in the source. The points P_1 and P_2 are perfectly coherent, although different atoms in the source are incoherent. This is known as spatial coherence.

If we now extend the size of the conventional source so that the path differences take on values which are no longer small compared to the wavelength, the vibrations at P_1, due to the source are no longer the same as at P_2. Some of the light at P_1 is coherent with some of the light of P_2. Therefore P_1 and P_2 are not perfectly coherent but neither are they incoherent. There is a partial coherence between P_1 and P_2 in spatial terms. As the source size increases then the partial spatial coherence decreases. Suppose we have two beams crossing, then for the incoherent case the intensity is given by :

$$I = A_1^2 + A_2^2 = I_1 + I_2$$

A_1A_2, I_1I_2 being the amplitudes and intensities of the two beams. If the two beams are perfectly coherent then the resulting intensity is given by :

$$I = A_1^2 + A_2^2 + 2A_1 A_2 \cos\phi = I_1 + I_2 + 2(I_1I_2)^{\frac{1}{2}}\cos\phi \quad .$$

For the two beams being neither perfectly coherent nor totally incoherent the intensity is :

$$I = I_1 + I_2 + |\Gamma_{21}|(I_1I_2)^{1/2}\cos\phi \qquad 3.3$$

where Γ_{21} describes the degree of coherence between the two beams. A more complete treatment can be found in Reference 8, Chapter X.

$|\Gamma_{21}|$ = 0 beams incoherent

= 1 beams perfectly coherent

Suppose we have an element of source dS, (Fig. 2) which produces an amplitude u_1 at P_1 and u_2 at P_2. We specify f_1 and f_2 as the complex transmission factors for P_1 to P and P_2 and P. At P element dS produces $f_1u_1 + f_2u_2$. Disturbances reaching P from other elements of the source will be incoherent with this. The intensity at P due to dS is given by :

$$dI_p = |f_1u_1 + f_2u_2|^2 \, dS$$

therefore

$$I_p = \int_s |f_1u_1 + f_2u_2|^2 \, dS$$

$$= |f_1|^2 \int_s |u_1|^2 \, dS + |f_2|^2 \int_s |u_2|^2 \, dS + 2 R_e \left\{ \int_s u_1^* u_2 f_1^* f_2 \, dS \right\}$$

$$= |f_1|^2 I_1 + |f_2|^2 I_2 + 2 (I_1 I_2)^{\frac{1}{2}} R_e (\Gamma_{21} f_1^* f_2) \qquad 3.4$$

where $\Gamma_{21} = \frac{1}{(I_1 I_2)^{\frac{1}{2}}} \int_s u_1^* u_2 \, dS \qquad 3.5$

Γ_{21} being the complex degree of coherence between P_1 and P_2. $V_{21} = |\Gamma_{21}|$ is the degree of coherence and $\beta_{21} = \arg\Gamma_{21}$ which is the phase difference of P_2 relative to P_1.

$$|f_1| = \frac{\text{real amplitude at P}}{\text{real amplitude at } P_1}$$

let $I_1' = |f_1|^2 I_1$ and $I_2' = |f_2|^2 I_2$, then 3.4 becomes :

$$I_p = I_1' + I_2' + 2 (I_1' I_2')^{\frac{1}{2}} R_e \left\{ \Gamma_{21} \left(\frac{I_1 I_2}{I_1' I_2'} \right)^{\frac{1}{2}} f_1^* f_2 \right\}$$

let arg f_1 = phase P - phase $P_1 = \phi_1$ similarly for arg f_2.

Then

$$I_p = I_1' + I_2' + 2 (I_1' I_2')^{1/2} R_e \left\{ \frac{\Gamma_{21} |f_1||f_2|}{|f_1||f_2|} e^{i(\phi_2 - \phi_1)} \right\}$$

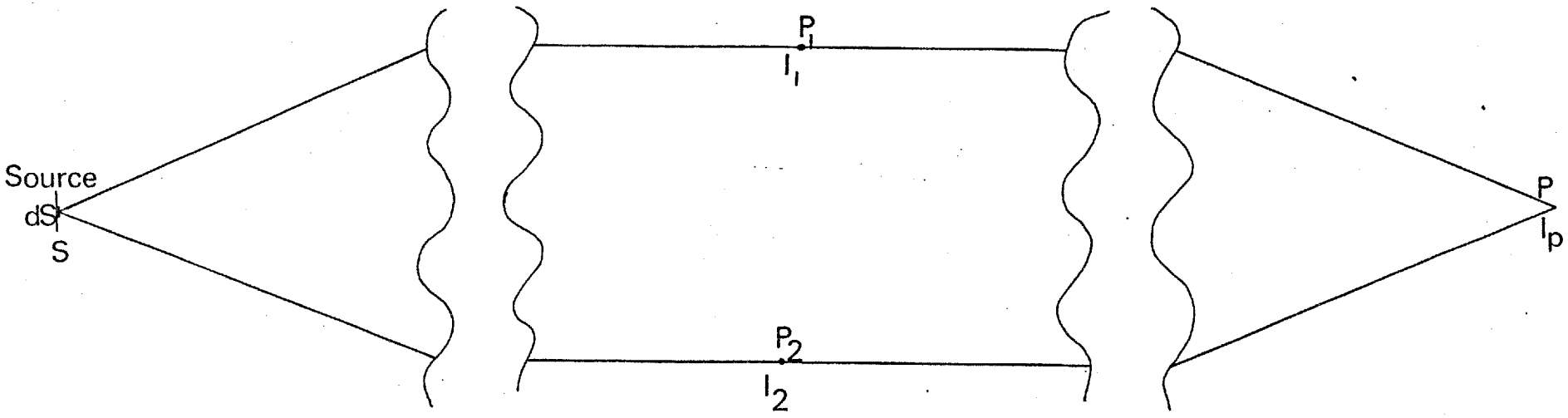


Fig 2 Coherence

therefore $$I_p = I_1' + I_2' + 2(I_1' I_2')^{1/2} R_e \{\Gamma_{21} e^{i(\phi_2 - \phi_1)}\} \qquad 3.6$$

now $$\Gamma_{21} = V_{21} e^{i\beta_{21}}$$

therefore $$I_p = I_1' + I_2' + 2V_{21}(I_1' I_2')^{1/2} \cos[(\phi_2 - \phi_1) + \beta_{21}] \qquad 3.7$$

Suppose we just have two holes at P_1 and P_2 then we will see interference fringes at P with :

$$I \begin{matrix} \max \\ \min \end{matrix} = I_1' + I_2' \pm 2 V_{21} (I_1' I_2')^{\frac{1}{2}}$$

assuming $$\phi_2 = 2m\pi\phi_1 \text{ and } \beta_{21} = 1 \quad .$$

The visibility or contrast of the fringes is given by :

$$\frac{I \max - I \min}{I \max + I \min} = \frac{2V_{21} (I_1' I_2')^{\frac{1}{2}}}{I_1' + I_2'} \quad . \qquad 3.8$$

From Equation 3.8 we see that the highest contrast fringes are obtained when $I_1' = I_2'$ the two beam intensities are the same and $V_{21} = 1$ we have perfect coherence between the interfering beams. For two beam interferometers such as the Mach-Zehnder P_1 travels along one arm of the interferometer and P_2 along the other arm, the contrast is as given by 3.8. The main factor which influences the contrast of the fringes is the source size as this determines the degree of coherence of the source.

Using Equation 3.5 it can be shown[8] that for a circular source of radius σ the degree of coherence

$$V_{21} = \left| \frac{2J_1\left(\frac{2\pi\sigma}{\lambda R}\rho\right)}{\left(\frac{2\pi\sigma}{\lambda R}\right)} \right|$$

where R is the distance to the plane containing P_1, P_2, the distance between P_1 P_2 is given by ρ and J_1 is the first order Bessel function. As the source size tends to zero the degree of coherence tends to unity as in the case of a point source. Good coherence is obtained when

$$\frac{2\pi\sigma}{\lambda R} \rho \leq 1$$

therefore $\rho \leq \frac{0.16\lambda}{\sigma/R}$. Now σ/R is the angular radius of the source which is given by sin α, α being the semi angle subtended at the position in question by the source. It then follows that the diameter of the circular area that is almost illuminated coherently by a quasi-monochromatic source is $\frac{0.16\lambda}{\sin \alpha}$. A quasi-monochromatic source is one whose light consists of spectral components that cover a frequency range $\Delta\nu$ which is small compared to the mean frequency ν. This result is used to estimate the maximum source size possible for good coherence when using a conventional source.

Complete incoherence occurs when :

$$\frac{2\pi\sigma}{\lambda R}\rho = 3.83$$

that is when the distance between $P_1P_2 = \frac{0.61\lambda}{\sin \alpha}$. Any source can be regarded as quasi-monochromatic for the Mach-Zehnder, since it is easy to compensate the paths.

III.2 Fringe Localisation

Consider a quasi-monochromatic point source S and two rays travelling along different paths meeting at S' where they interfere, (Fig. 3). The phase difference at S' is :

$$\delta = \frac{2\pi}{\lambda}\left\{[SA_2\theta_2S'] - [SA_1\theta_1S']\right\} \qquad 3.9$$

The value of δ depends upon the position of S' but is uniquely determined for all S'. The interference fringes follow the loci of points for which δ is constant and can be formed at any plane where the two paths from S intersect. These fringes, from a point source, are not localised.

Let us now consider an extended source at S and let M be some point in this source (Fig. 4). The visibility of the fringes at some position of S' remains at or near that of a point source, elsewhere it has fallen well below. The fringes are said to be localised and occur when an

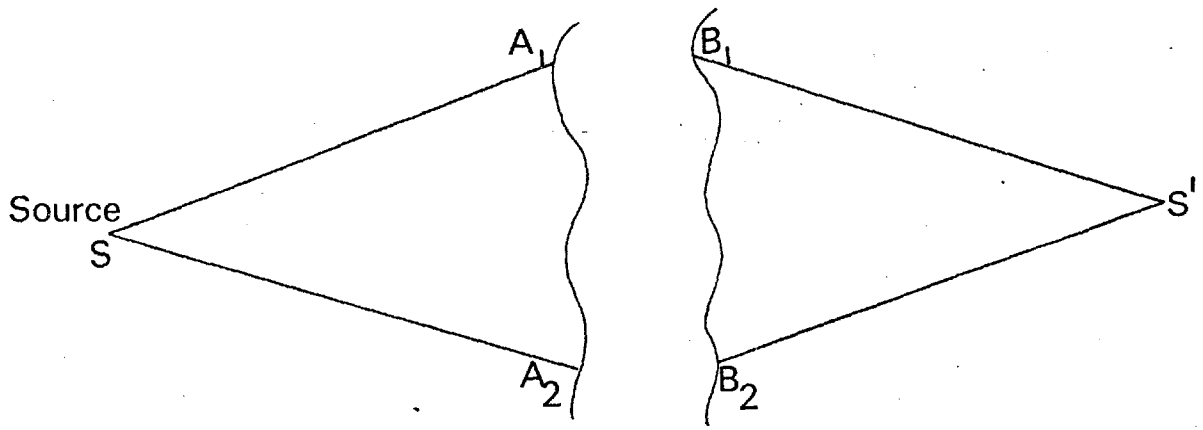


Fig 3 Fringe Localisation

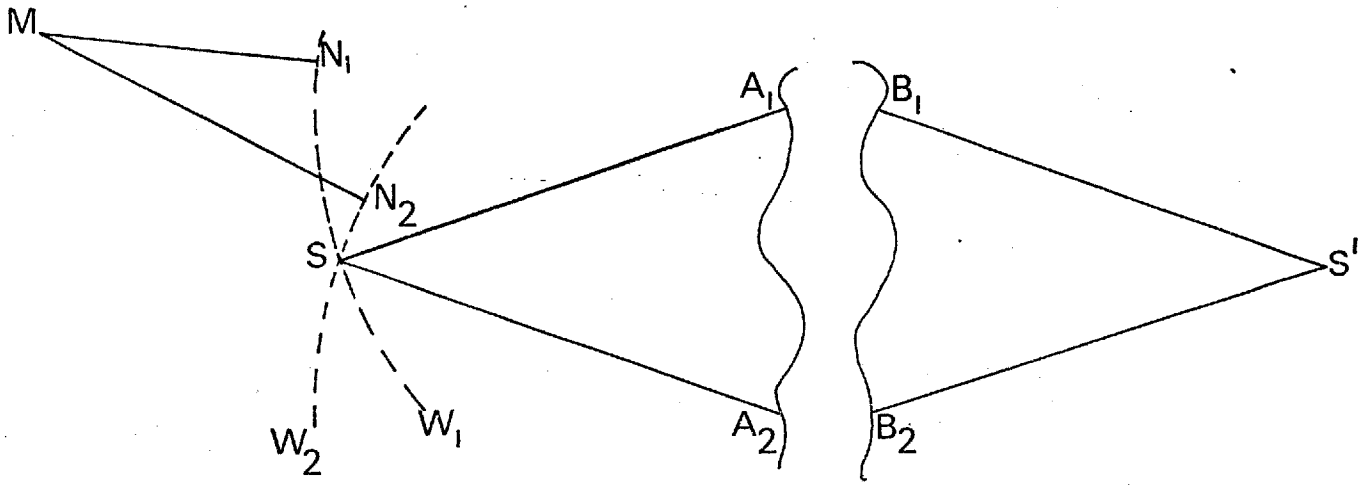


Fig 4 Path Lengths

extended source is used. We wish to know where to focus. To do this we reverse the light direction and take wavefronts back from S' and see what happens at S.

W_1 and W_2 are wavefronts, which are two spheres with radii R_1 and R_2, normal to SA_1 and SA_2. MN_1 and MN_2 are normals to these wavefronts. The optical paths from N_1 to S' and N_2 to S' are $[SA_1B_1S']$ and $[SA_2B_2S']$ respectively assuming both paths are in air with refractive index = 1. The path difference at S' between the two rays of M is :

$$\delta_m = [MN_2] + [SA_2B_2S'] - [MN_1] - [SA_1B_1S']$$

from 3.9 using δ to represent path difference.

$$\delta m - \delta = l_2 - l_1 \qquad 3.10$$

where $l_2 = [MN_2]$ and $l_1 = [MN_1]$.

To evaluate 3.10 we take S to be the origin of a rectangular co-ordinate system SX, SY, SZ, where SX is the internal bisector and SZ the external bisector of angle A_1SA_2, (Fig. 5). If we assume that point M is in the neighbourhood of S it can be shown from simple geometry that :

$$\delta_m - \delta = y\theta + \tfrac{1}{2}\left(\frac{1}{R_2} - \frac{1}{R_1}\right)(y^2+z^2) \qquad 3.11$$

Equation 3.11 gives us the path difference in S' (Fig. 4) when passing from point S to point M in the source. If we choose S' in a position such that the rays from the source coincide, i.e. the wavefronts W_1 and W_2 are tangential, then the linear term $y\theta$ in 3.11 vanishes. Using this we can obtain a criterion for the largest permissible source size. This is that in order to have good visibility fringes at S' we must have $\delta m - \delta$ smaller than a quarter of a wavelength. Alternatively, for a given source size the fringes have maximum visibility at this point.

For a particular source size, the interference fringes are localised

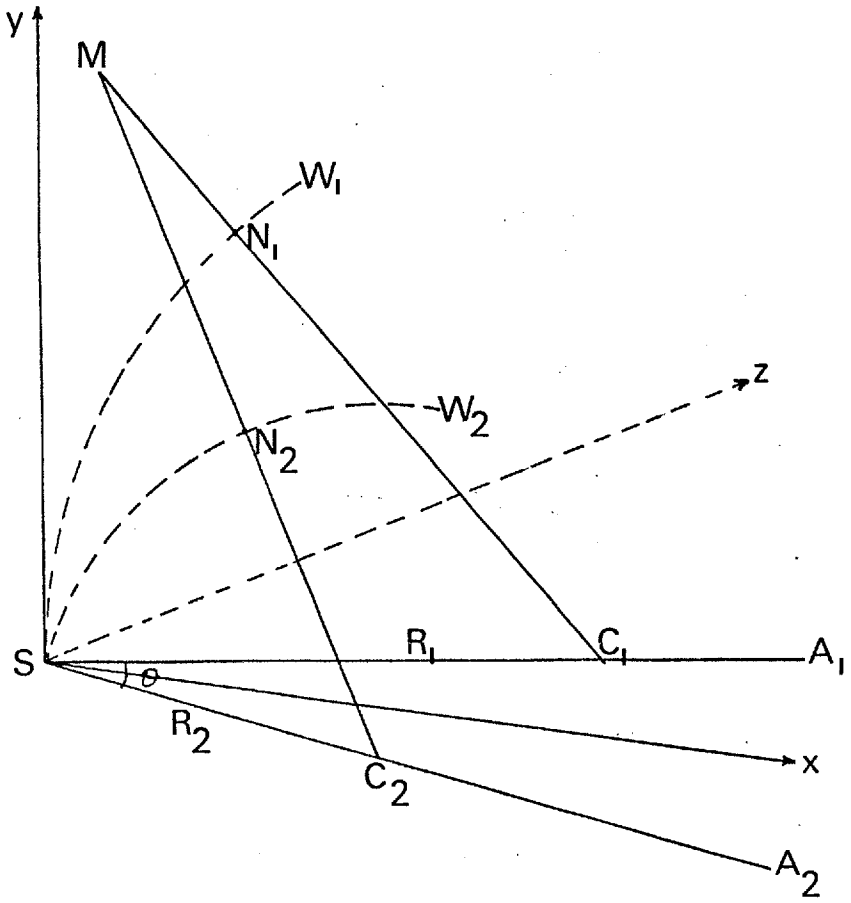


Fig 5 Path Differences

on a surface which is the locus of points of the intersecting rays originating from the source. The region of localisation is the region where the two rays cross. In the Mach-Zehnder interferometer we have the facility to be able to vary the position of localisation so that it can coincide with the position of the test sample. Photographs of the fringes occuring in this plane can then be taken.

III.3 The Interferometer

The Mach-Zehnder interferometer, shown schematically in Fig.6, operates with the division of amplitude. A full account of this and other similar types of interferometers can be found in Refs.[8,13,21,23,74,79].

The Mach-Zehnder interferometer consists basically of two mirrors and two beamsplitters. With adjustments being possible on all four of these optical components its initialy setting-up is a little more lengthy than other similar interferometers such as the Twyman-Green. On the other hand it is more versatile since the difference between the two beams are fairly generally adjustable. Its sensitivity, although less than double path and multiple beam interferometers, is sufficient for our present use. Its main features are that is has the facility to vary the position of localisation of the fringes, and also to be able to equalise the optical path lengths of the two beams. These two features assist in providing good contrast fringes with a reasonable amount of illumination.

The Mach-Zehnder interferometer built for the present study consists of two beamsplitters coated for approximately 50% transmission and 50% reflection over a neutral spectral range, and two high reflectance mirrors. Each beamsplitter is wedged to about 1^{o} and the mirrors are front surface aluminised. This is to avoid inhomogeneities in

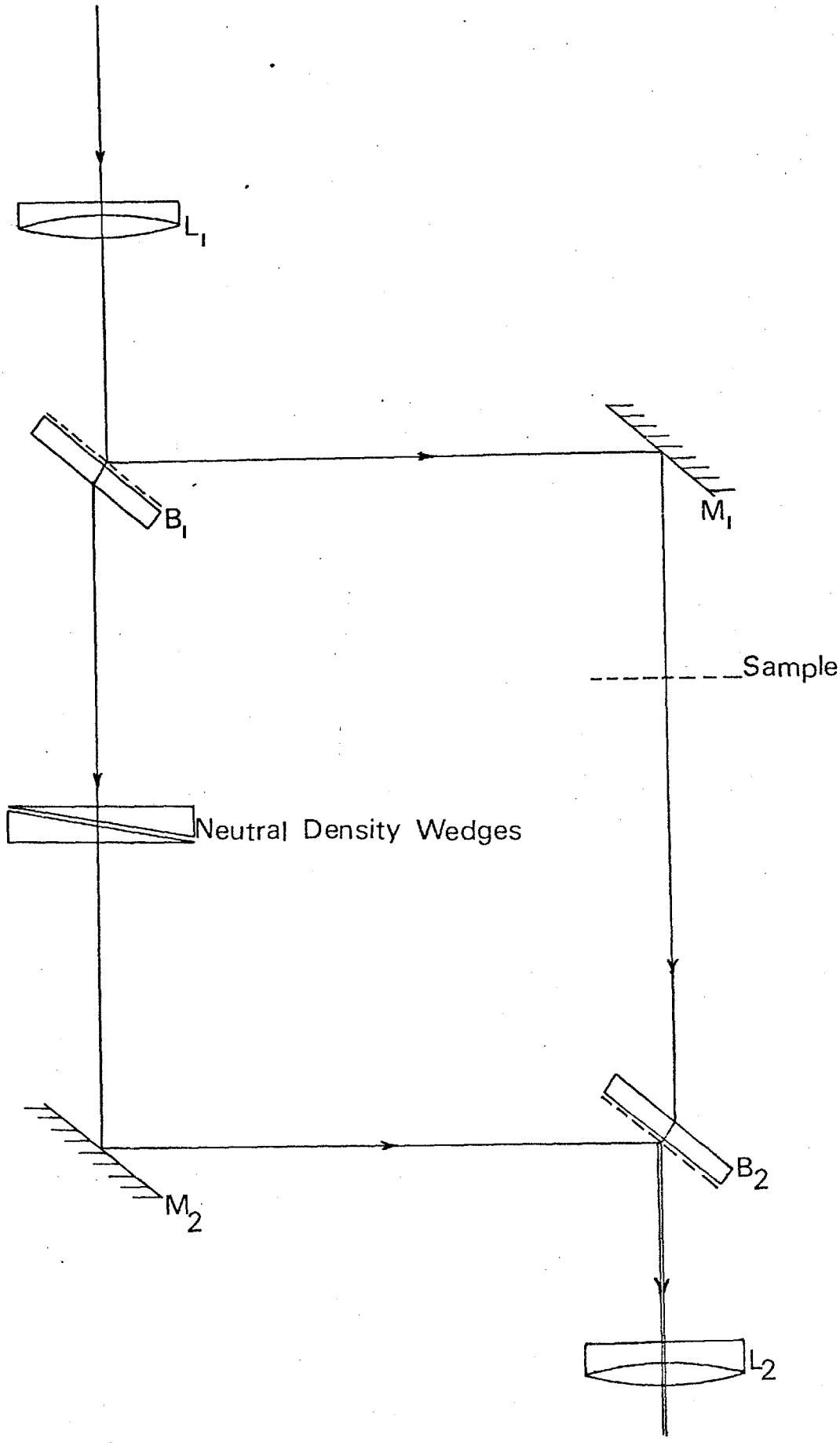


Fig 6 Mach-Zehnder Interferometer

the glass as well as avoiding multiple beam interference fringes. All plane surfaces in the interferometer are polished flat to $\frac{1}{10}$ of a wavelength. The dimensions of the mirrors and beamsplitters are 6 cm x 6 cm giving a maximum inspection diameter of the order of 4 cm which is adequate for examining 35 mm film. The interferometer was set up on a two foot square steel table.

From Equation 3.8 it is seen that the best contrast fringes are obtained when the intensities of the two interfering beams are the same. As some of the emulsions being studies could have a density of over two, this would result in a serious loss of contrast (for $I_1 : I_2$ = 100:1 contrast = 0.2). In order to compensate for the different densities of emulsion two neutral density wedges, flat to one wavelength in transmission over the whole surface, were constructed and placed in the reference arm of the interferometer. The dimensions and tolerances for these wedges are shown in Fig. 7. These two wedges were made from Schott NG 11 glass (the transmission curves for this glass can be seen in Fig. 8) and mounted on precision low friction slides. These slides, shown schematically in Fig. 9, enabled the wedges to be racked across the beam, thus varying the intensity transmitted in that beam, without introducing any random tilt to the wedges. One of the wedges was tilted by about 1^{o} so as to avoid multiple beam interference fringes between the wedges. Using these wedges sample densities of up to three can be examined without loss of contrast due to unfavourable beam intensities. The use of the wedges does however introduce a small path length change in that arm of the interferometer which would reduce contrast. By initial careful setting-up with the wedges in position it was found that there was no serious loss of contrast. The interferometer was

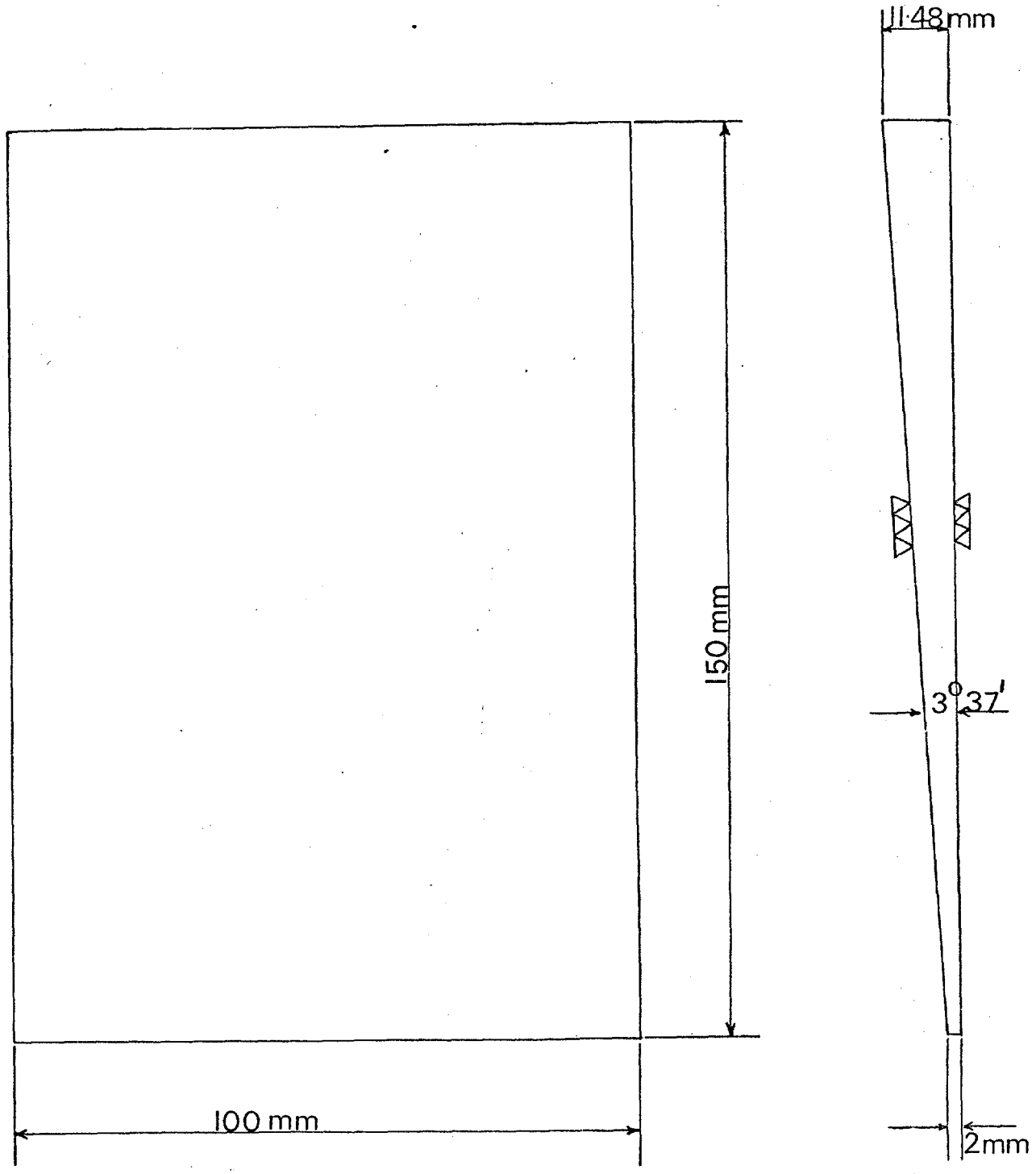


Fig 7 Neutral Density Wedge Flat to One Wavelength

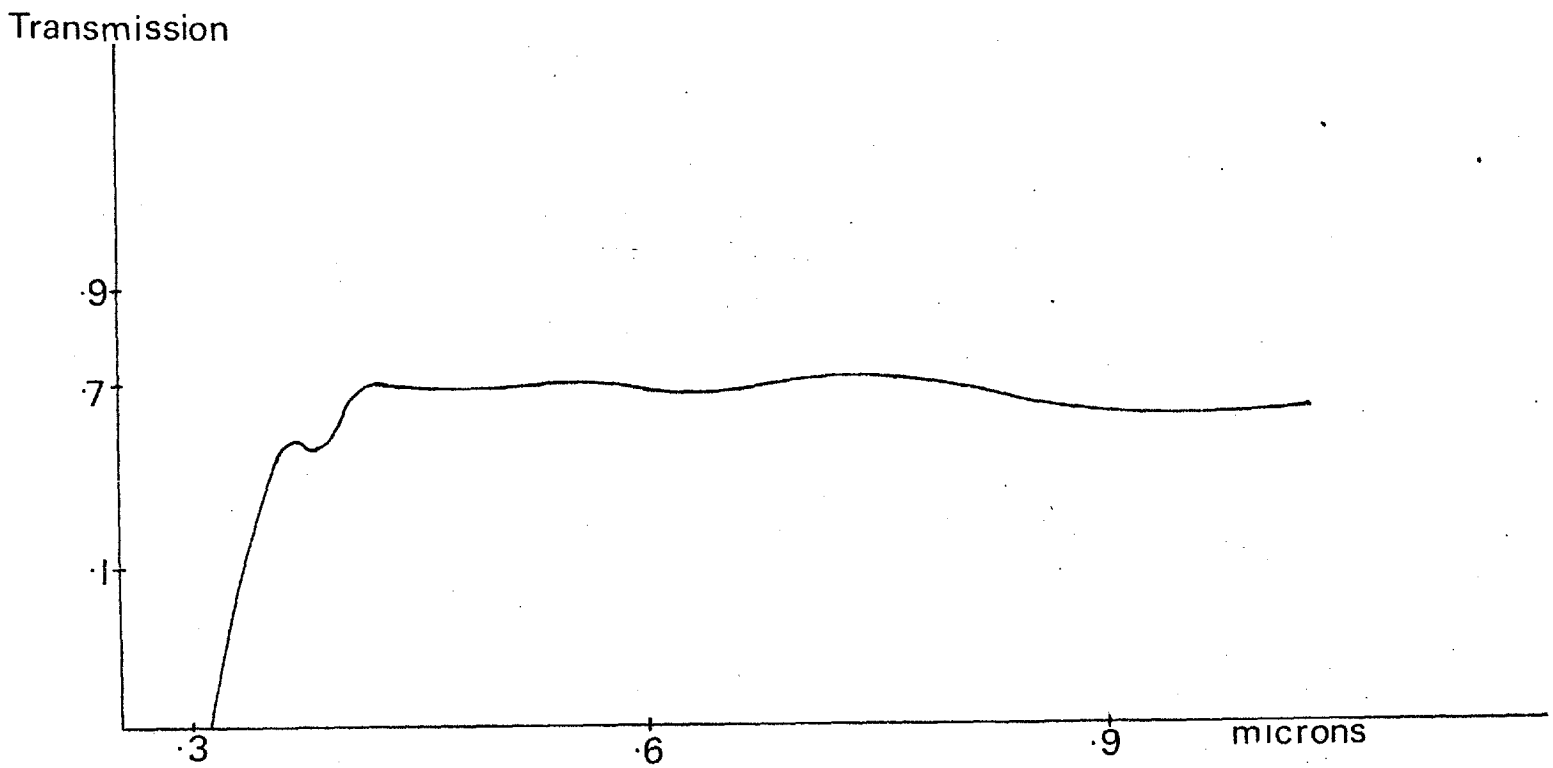
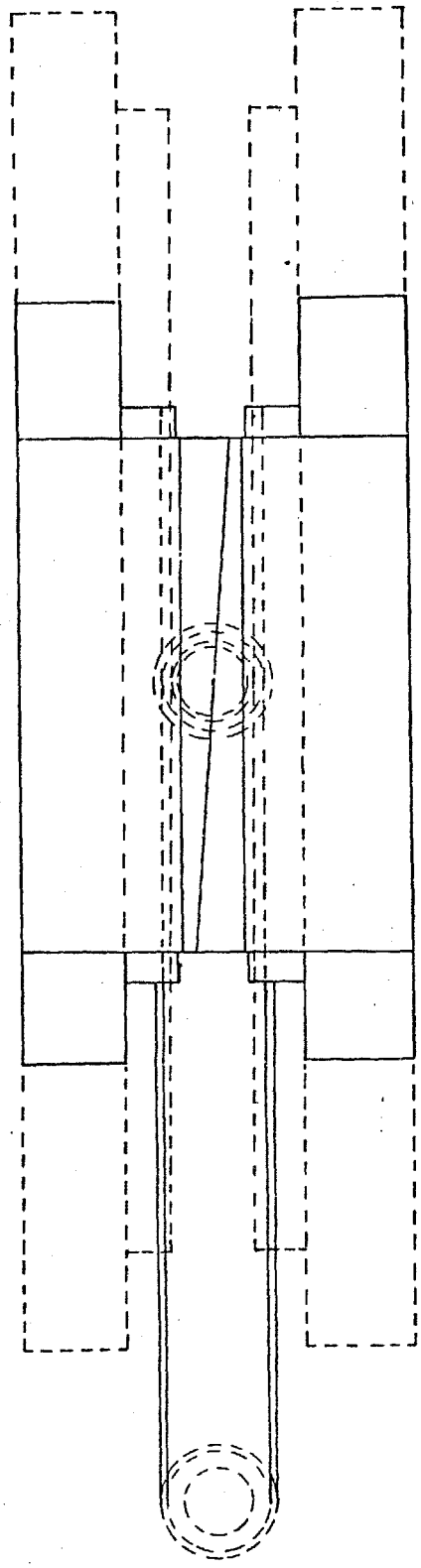
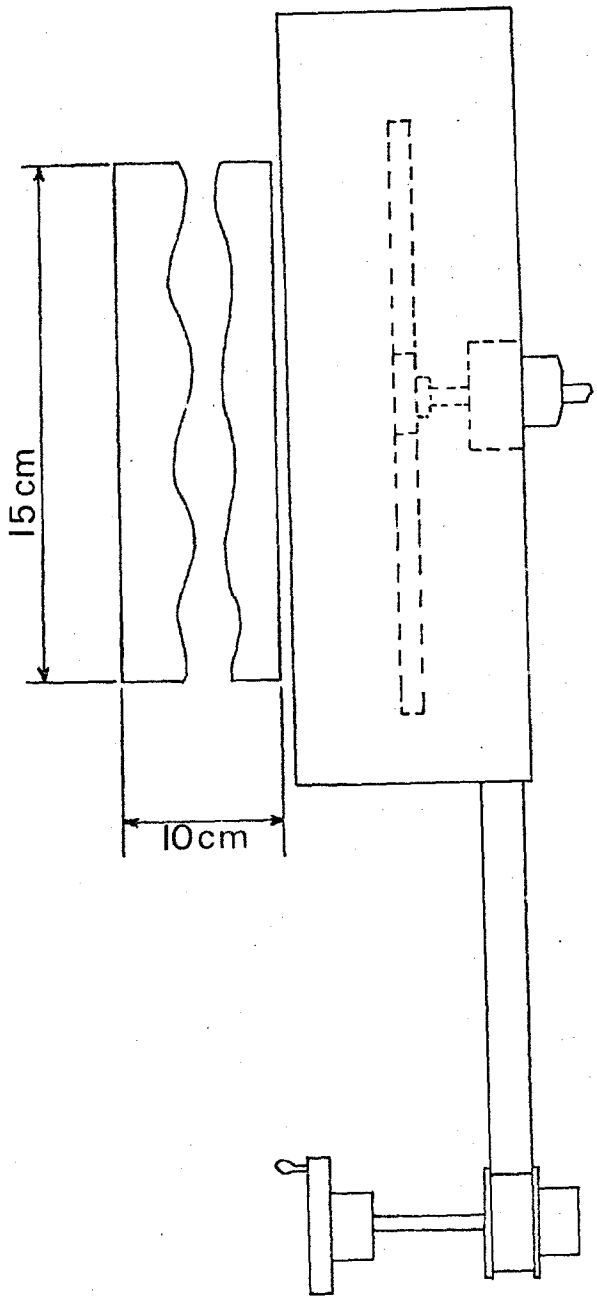


Fig 8 Transmission Curve for 1mm NG.11





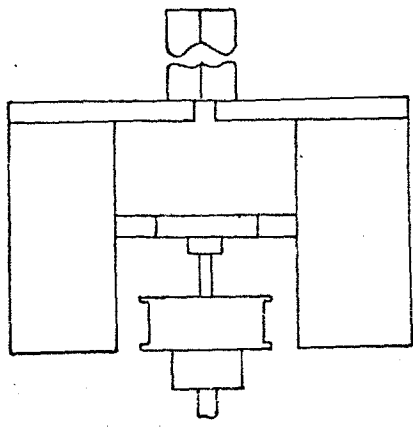


Fig 9 Slides for Neutral Density Wedges. Drawn 1/2 scale

initially set-up using laser illumination, thus easily obtaining good quality fringes. The laser was then replaced by a conventional source, a low pressure mercury lamp, and the position of the mirrors and beam-splitters was adjusted until the highest contrast fringes were obtained. The adjustment was carried out by trying a tilt in one direction and then observing the fringes to see whether they were clearer or not when they had been re-fluffed out. The process was repeated in small stages until the best contrast fringes were obtained. It was assumed that this position corresponded to the theoretical best setting of the interferometer.

It was found that the clarity of the fringes was better when using the low pressure mercury lamp rather than the laser source. This was due to the speckle and diffraction effects of dust and other such unwanted particles which occured with the laser illumination. However, when sample densities of about 1.5 and over were being examined, the mercury lamp was not of sufficient intensity to produce fringes which could easily be photographed. Increasing the source size only reduced contrast. For higher densities therefore, a laser source was used. All results quoted in units of λ in the following chapters refer to the wavelength of the laser line of 6328Å. All results using the mercury source have been converted to the laser wavelength.

The mounts for the mirrors and beamsplitters have been made following the basic design of the mounts used in a Hilger and Watts Twyman Green interferometer. They give tilt movement in a horizontal and vertical direction pivoting about the base of the mount. This is shown in Figs. 10 and 11. Fine adjustment is achieved by having the movement geared down by a lever arm to the base. Using such a

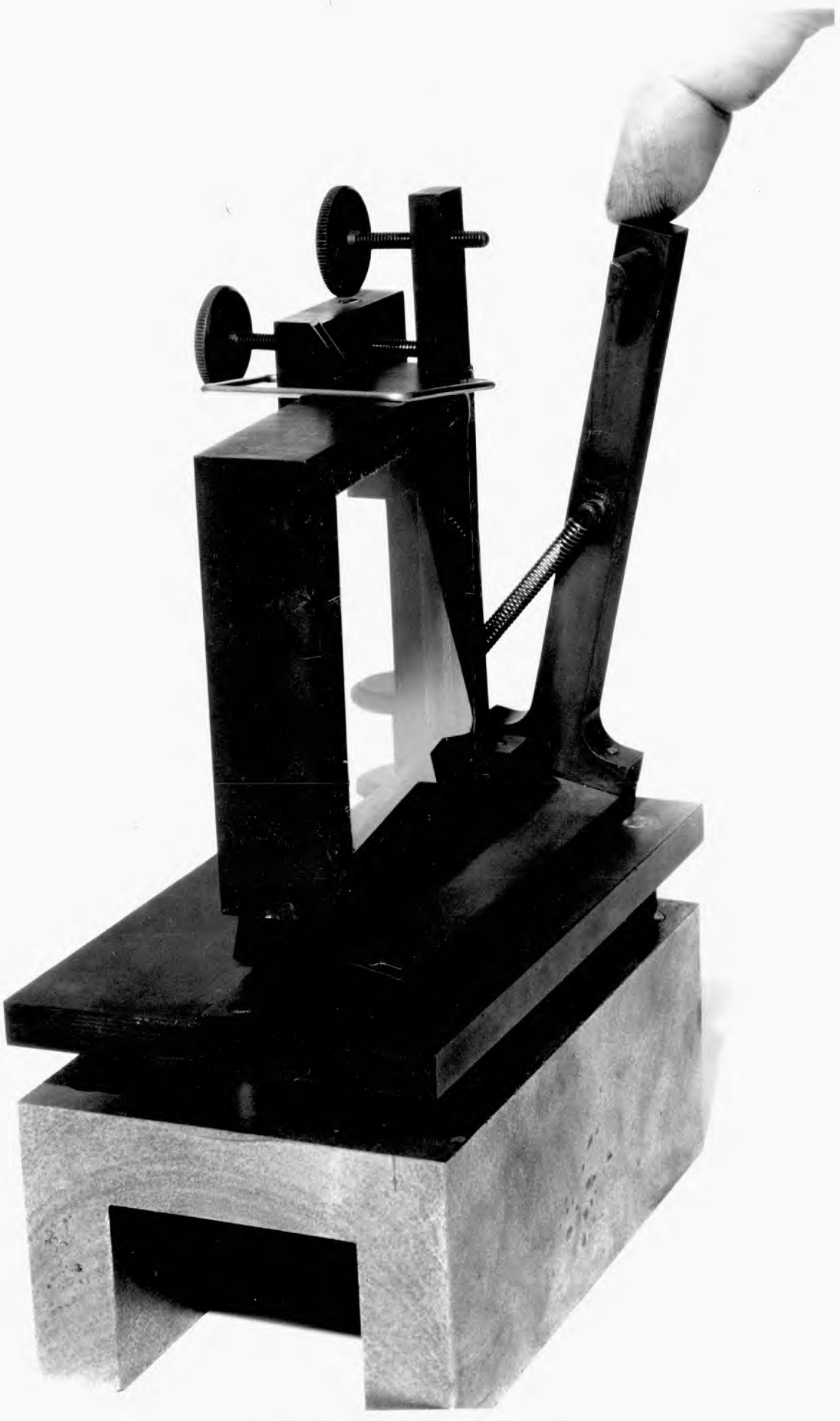


Fig 10 Horizontal movement

Fig 11 Tilt movement

lever arm it was found that one complete screw turn gave a tilt of four minutes. Ideally these mounts should have been designed so as to pivot about the centre of the beam, thus avoiding path length change when tilting. This effect is marginal and only really essential if a white light source or source having a short coherence length is used. In any case, once the inteferometer is set-up no adjustment is made during the course of examining the various emulsions for a particular set of results.

Light from a source S (Fig. 6), either a low pressure mercury lamp or a laser focused down onto a limiting aperture, is collimated by lens L_1 producing a parallel beam of light of about 3.5 cm diameter incident onto beamsplitter B_1, The two beams are then reflected by mirrors M_1 and M_2 and made to recombine at the semi-reflecting surface of an identical beamsplitter B_2. The emergent beam is then brought to a focus by lens L_2. The mirrors and beamsplitters are set-up in a rectangular array with their optical surfaces parallel. When this ideal set-up is achieved the two plane wavefronts recombining at B_2 would be exactly in phase and would therefore interfere constructively to produce a bright field. If now either of the mirrors is tilted, (Fig. 12), the two plane wavefronts W_1 and W_2 would be inclined with respect to each other with an air wedge between them. W_1' is the virtual image of wavefront W_1 and in phase with it. At a point P on W_2 the phase difference between this and the point H on W_1' normal to it is :

$$\delta = \frac{2\pi}{\lambda} d$$

where d is the distance between P and H. A bright fringe will be observed when the path difference

$$d = m\lambda \ , \quad |m| = 0, 1, 2 \ldots$$

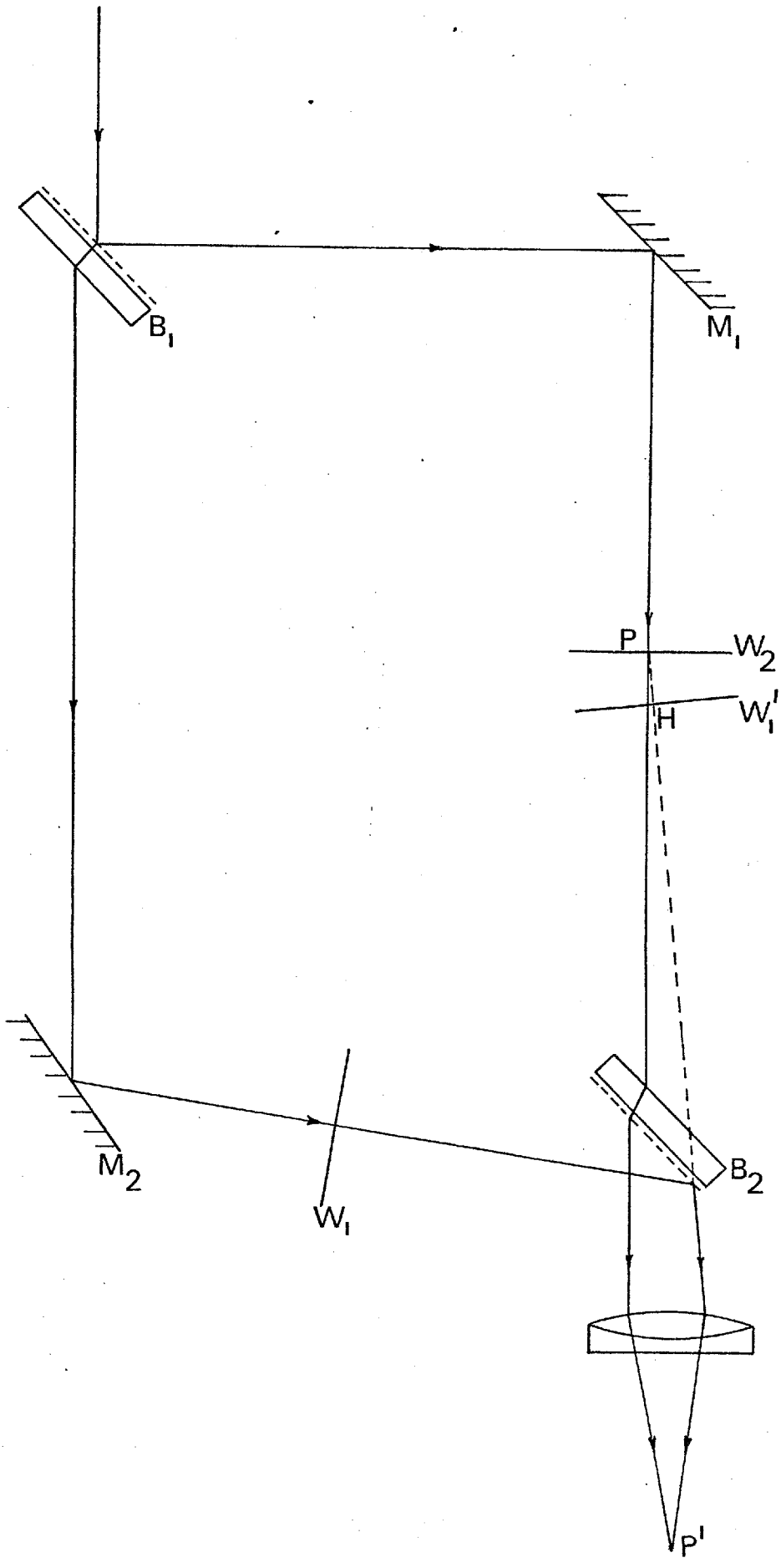


Fig 12 Fringe Localisation

and a dark fringe is seen when :

$$d = n\lambda \ , \ |n| = \frac{1}{2}, \frac{3}{2} , \frac{5}{2} \ \ldots.$$

i.e. the separation of dark to dark or bright to bright fringe is the result of one wavelength of path difference being introduced.

In normal use the interferometer is set with the two wavefronts inclined with respect to each other and using the largest source size possible for good contrast fringes. In this configuration, straight line fringes are observed. The position at which the two "virtual" wavefronts cross is the position of localisation of the fringes and it is in this position that the highest contrast fringes are observed. The position of localisation can be adjusted by a combination of rotation of the mirrors and beamsplitter B_2 to coincide with the plane of the test sample, and it is this plane which is the object plane of the lens L_2. If initially all the optical surfaces are parallel then the two wavefronts are parallel and the fringes are localised at infinity. If mirror M_2 is then rotated slightly the region of localisation lies close to M_2. By rotating beamsplitter B_2 we can move the region of localisation towards B_2. A similar change of localisation can be achieved by rotating mirror M_1.

From examination of Equation 3.11 we see that the first term gives us the region of localisation when using an extended source and the second term gives us the limitations on source size. If $\theta = 0$ and we assume a circular source centered on S (Fig. 5), Equation 3.11 reduces to :

$$\delta m - \delta = \tfrac{1}{2} \left(\frac{1}{R_2} - \frac{1}{R_1}\right) r^2 \qquad 3.12$$

where $y^2 + z^2 = r^2$, r being the radius of the source.

For good visibility fringes we must have the path difference smaller than a quarter of a wavelength. Then the radius of the source is given by :

$$r \leq \left(\frac{\lambda R_2 R_1}{2|R_2 - R_1|} \right)^{1/2} \qquad 3.13$$

For the special case of $R_2 = R_1$ and the paths along each arm of the interferometer being equal, we can, theoretically, observe good visibility fringes with any source size. The fringes of the Mach-Zehnder interferometer are formed from an air wedge with the source at infinity. Let 2α be the angle the source subtends at the collimating lens. Then this gives us that the angle subtended by the source at the collimating lens as :

$$\alpha \leq \left(\frac{\lambda}{2|R_2 - R_1|} \right)^{1/2}$$

which is the limitation on source size for good visibility fringes. As the distance between the wavefront decreases the source size can increase without loss of contrast.

In practice we have a secondary consideration on source size, namely that of light intensity. In practice therefore we use that source size which gives us the best contrast fringes together with a useable light intensity. When using the mercury lamp it was found that the largest usable source size subtended an angle of approximately 0.007 degrees at the collimating lens.

The fringes obtained using the interferometer were photographed onto 35 mm film using a telephoto lens.

CHAPTER IV - THE COHERENT PROCESSOR

When a recording medium such as a photographic emulsion is used in a coherent optical system we need to know its complex transmission properties. The phase properties are found by interferometric means as described in the previous chapter. In order to investigate the amplitude transmission properties we can use a coherent optical processor. In normal use the coherent processor is used to perform operations on certain signals recorded on the emulsion. The same operations are of course also performed on the emulsion. The resultant effect is the required output signal together with a noise contribution from the emulsion. If we were to perform the operations using a piece of emulsion which contains no recorded signal then the output would be a function of the complex transmission properties of the emulsion. It must be realised that the optical elements forming the processor are themselves not perfect and would produce a degrading effect on the output. These contributions can be accounted for by evaluating the performance characteristics of the system.

IV.1 Diffraction Properties of a Circular Aperture

For reasons of circular symmetry we will use a circular aperture to limit the system. Let the co-ordinates of a typical point in the aperture be $(\rho\theta)$ and that of a point P in the diffraction pattern be $(\omega\psi)$ (Fig.1).

We have the following relationships

$$\xi = \rho\cos\theta \quad , \quad \eta = \rho\sin\theta$$

$$\rho = \omega\cos\psi \quad , \quad q = \omega\sin\psi$$

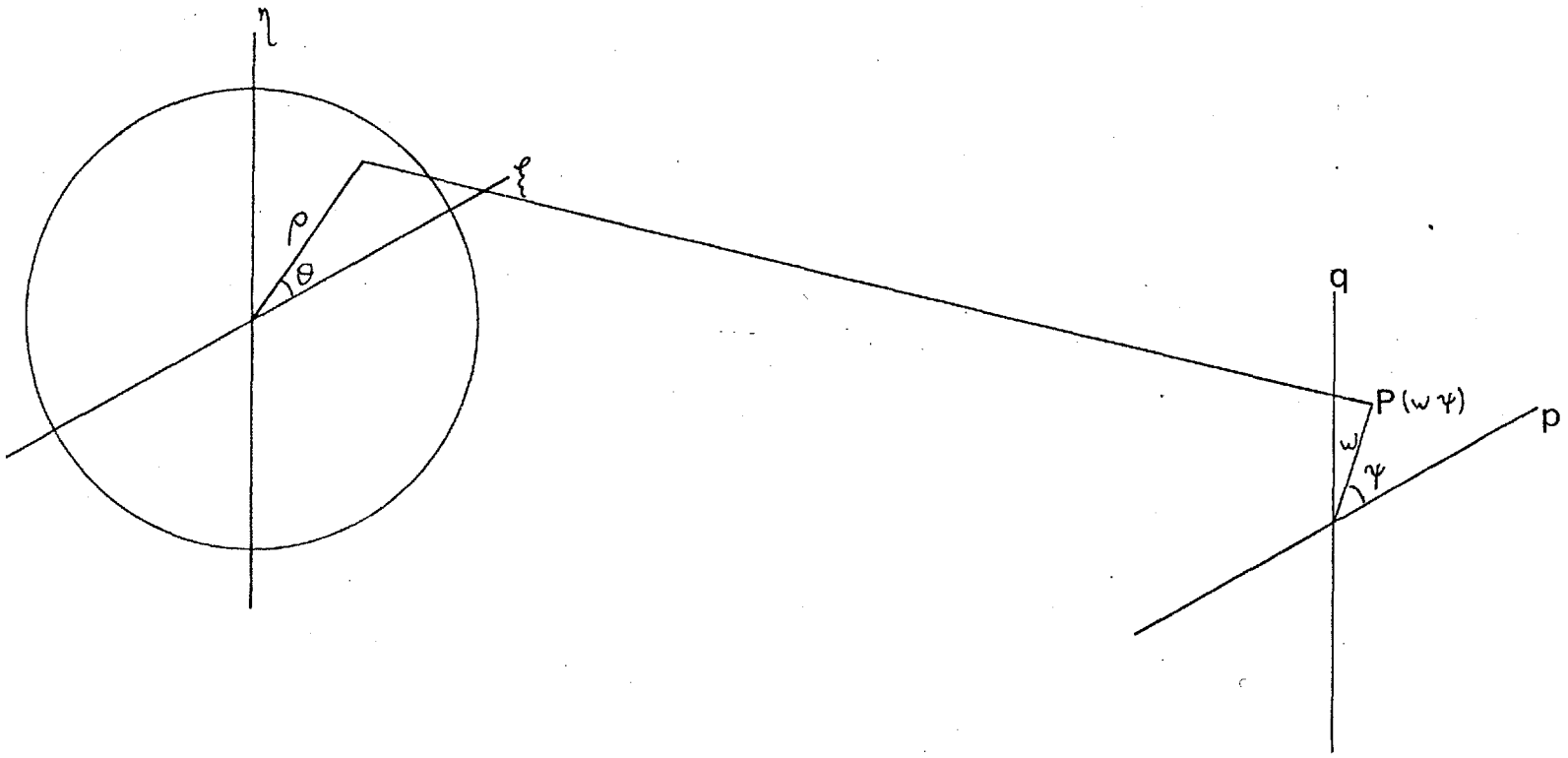


Fig 1 Co-ordinate System

The distribution of light intensity is the modulus of the square of the Fourier transform of the circular aperture, therefore the complex amplitude at P is given by :

$$U_\rho = K\int_0^a\int_0^{2\pi} e^{-ik\rho\omega\cos(\theta-\psi)}\,\rho\,d\rho\,d\theta \qquad 4.1$$

where a is the radius of the aperture and K is a constant

now $$\int_0^{2\pi} e^{iz\cos(\theta-\psi)}\,d\theta = 2\pi J_0(z)$$

where Jo is the zero order Bessel function. Equation 4.1 now becomes :

$$U_\rho = 2\pi K\int_0^a J_0(k\rho\omega)\,\rho\,d\rho \qquad 4.2$$

now $$\int_0^z z_0 J_0(z_0)\,dz_0 = z J_1(z)$$

let $z_0 = bt_0$, $$\int_0^{bt} bt_0 J_0(bt_0)\,b\,dt = bt J_1(bt)$$

therefore $$\int_0^{bt} t_0 J_0(bt_0) = \frac{t J_1(bt)}{b}$$

Using this relationship in 4.2 we get :

$$U_\rho = K'\left[\frac{2J_1(ka\omega)}{ka\omega}\right] \qquad 4.3$$

where $K' = \pi a^2 K$. The intensity in the diffraction plane is :

$$I_\rho = |U_\rho|^2 = I_0\left[\frac{2J_1(ka\omega)}{ka\omega}\right]^2 \qquad 4.4$$

which is known as the Airy disc. We are not so much interested in the actual light intensity but the distribution of light intensity in the diffraction plane. This is given by the first order Bessel function $\left[\frac{J_1(ka\omega)}{ka\omega}\right]^2$ and is seen in Fig. 2. The function of Fig. 2 has a maximum value at x = 0, which is the centre of the Airy disc. As we move away from the centre the function decreases until the intensity becomes zero at $J_1(x) = 0$ after which it oscillates with ever decreasing amplitudes. The position and amplitudes of the maxima and minima can be evaluated.

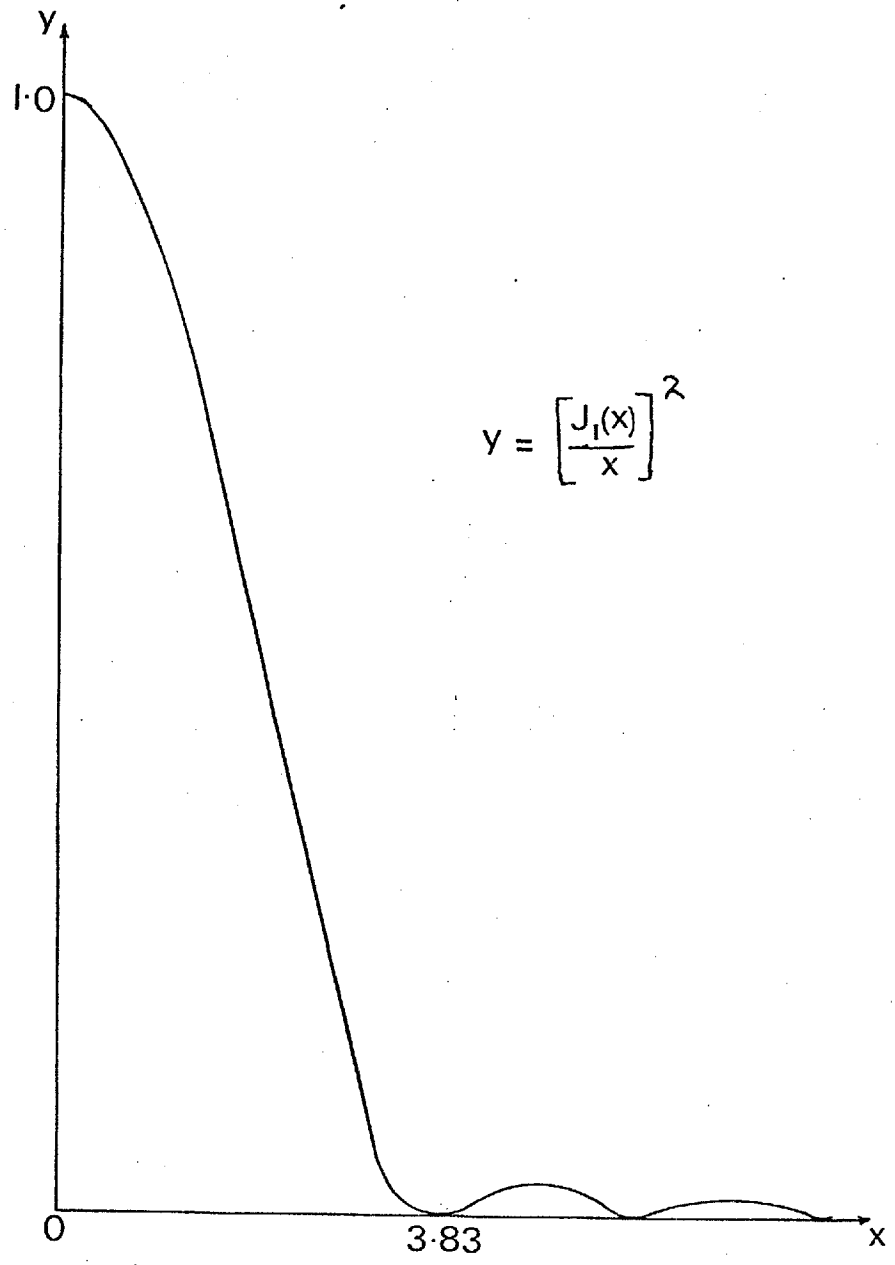


Fig 2 Airy Disc

Zero intensities occur at all positions of x satisfying $J_1(x) = 0$ which can be found from tables. Similarly the maxima occur at values of x satisfying

$$\frac{d}{dx}\left[\frac{J_1(x)}{x}\right] = 0$$

The resultant pattern consists of a bright central portion centered on the geometrical image p = q = 0 of the source, surrounded by a series of concentric bright and dark rings. The intensity of the bright rings decreases with their distance from the centre.

In coherent processing we are really dealing with amplitude spread functions. The term $\frac{2J_1(ka\omega)}{ka\omega}$ is the Fourier transform of the circular aperture and $\left[\frac{2J_1(ka\omega)}{ka\omega}\right]^2$ is its power spectrum.

To complete the picture we will consider briefly the phase distribution near the focus (Ref. 8, Sec. 8.8.4). Here it is shown that the surfaces of equal phase are substantially plane close to the focus. However, the intensity across these phase surfaces is not uniform. In the immediate neighbourhood of the Airy disc dark rings these equal phase surfaces become more complicated. Going away from the geometrical focus along the focal plane the phase is constant between any two successive dark rings, but changes abruptly by π on crossing each ring. The same phase jump occurs on crossing each point of zero intensity along optic axis.

If a photographic emulsion is placed in the plane of the circular diffracting aperture, the diffraction pattern so formed is modified from that of above and the resultant pattern is a representation of the emulsion. Therefore the complex transmission of the emulsion can be characterised by its influence on the diffraction pattern.

IV.2 The Coherent Processor

The coherent processor is shown in Fig. 3. Light from a laser

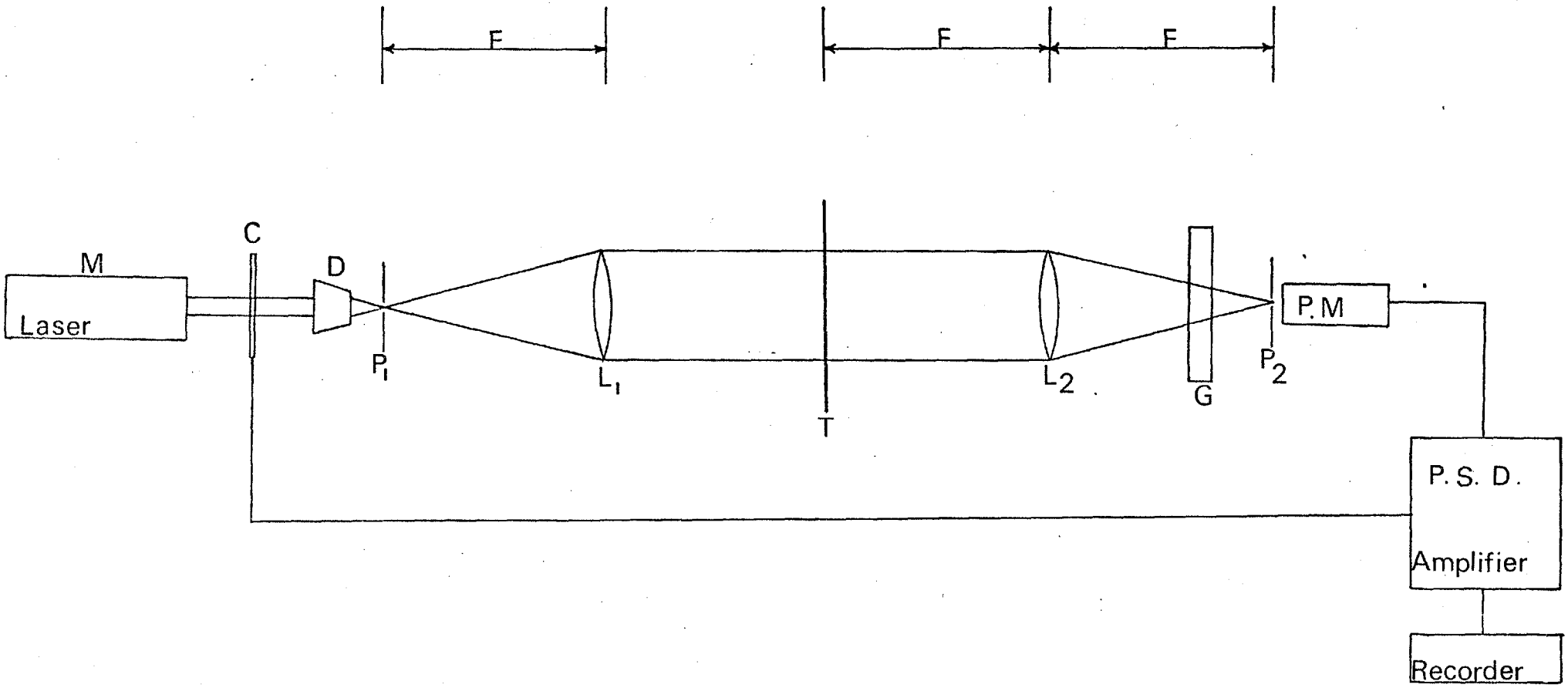


Fig 3 Coherent Processor

source, M, chopped by a variable frequency mechanical chopper C, is focused down by a microscope objective ϕ on to a pinhole P_1. The size of the pinhole used is dependent upon the numerical aperture of the microscope objective. The radius of the first Airy ring of the objective is given by $\frac{0.61\lambda}{NA}$. The diameter of the pinhole was chosen to be less than the diameter of the central intensity of the Airy disc. (The size of the pinhole was approximately checked with a microscope). This gave a reasonably clean beam, together with the maximum possible light intensity. The best combination for the system was found to be a x10 microscope objective with a five micron pinhole. Lens L_1 is used to collimate the emergent light beam which then passes through the transparency T. A second lens L_2 situated at the focal distance F from the transparency, displays the two dimensional Fourier transform of the complex amplitude transmission through the transparency, at a distance F on the image side. A second pinhole, P_2, is accurately located in the image plane of the lens L_2. The diameter of pinhole P_2 is chosen so that it is an order down on the radius of the first Airy ring produced by lens L_2. The pinhole used was found to give a well defined first minimum for the largest aperture used. Light which passes through this pinhole strikes the cathode of a photomultiplier P.M. The type of photomultiplier used was an E.M.I. 9592B having an S.10 photocathode. The quantum efficiency of this type of photocathode is of the order of 2%. The voltage supply to the photomultiplier was provided by a stabilised E.H.T. power supply. The output from the photomultiplier goes through a phase sensitive detector[10] and a switched gain amplifier and the signal is recorded on an x-y recorder.

The circuit diagram for the phase sensitive detector is shown in Fig. 4. The frequency of the chopper is tuned to match an electronic filter thus reducing noise level due to stray light. A reference signal is sent from the chopper to the phase sensitive detector, and the phase of this is compared to the phase of the signal from the photomultiplier. The bandwidth of the phase sensitive detector is given by $\frac{1}{2\pi RC}$ where R and C are the values of the resistance and capacitance on the output of the detector. For the present case this was of the order of 6 c/s. The relative phases of the two signals are adjusted by changing the position at which the laser beam passes through the chopper. The position of the chopper giving maximum output signal from the phase sensitive detector corresponds to the two signals being in phase, and it was this configuration which was used in recording the results.

In order to scan the image produced by lens L_2 a piece of plane parallel glass (polished flat to $\frac{1}{10}$ of a wavelength in transmission) is placed between lens L_2 and its image plane. The rotation of this glass is used to scan the image across the pinhole P_2. The speed of the scan can be varied by varying the voltage to the motor rotating the glass and this is synchronised with the output of the recorder. Over the limited range of rotation of the glass the rotation can be considered to be linearly related to distance. In any case any non-linearity is allowed for by normalising the distances to those obtained without using any emulsion. Thus we are able to provide an approximately linear relationship between distances in the Airy pattern located at the image plane of L_2 and the distances of the Airy pattern recorded on the x-y recorder. The axis of rotation of the glass was accurately defined in two directions at right angles

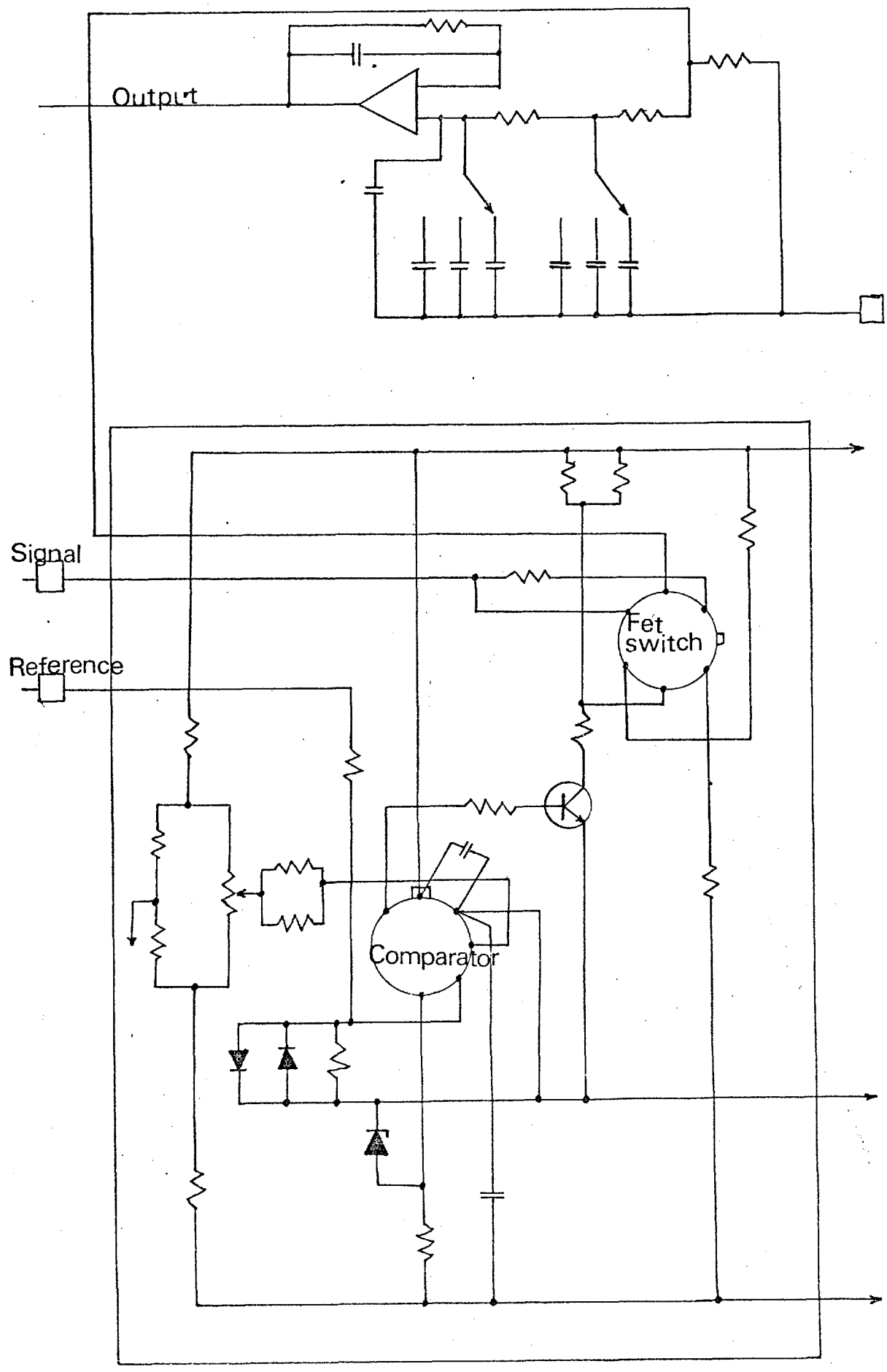


Fig 4 Phase Sensitive Detector. Filter

to each other and at 45^{o} between them. Scans can therefore be made in both radial and tangential azimuths or along and across the plane of the emulsion.

The linearity of the photomultiplier was checked by using a series of accurately known optical density gelatin filters. From this check it was found that the response of the photomultiplier was linear over a wide range of intensities for the E.H.T. voltages used. Slight non-linearity was observed for low light level signals. This was probably due to photon noise. The effect of this non-linearity was taken into account in all low intensity light signals.

Due to the random nature of the phase variations found experimentally, it was found necessary to incorporate the interferometer into the coherent processor, (Fig. 5), thus performing both experiments under identical conditions. We now make use of the fact that we can obtain two sets of fringes after beamsplitter B_2 as shown. As before we can photograph the interferograms by using lens L_3 and a suitable camera. When doing this we have of course, no need to chop the laser beam. In order to scan the diffraction pattern we set the mechanical chopper in motion and obstruct the reference beam of the interferometer by placing an opaque object somewhere between $B_1M_1B_2$. Lens L_2 and the system following it is then used in the same manner as the ordinary optical processor. Hence we are able to perform both the experiments on phase and scans of the Airy disc under the same conditions.

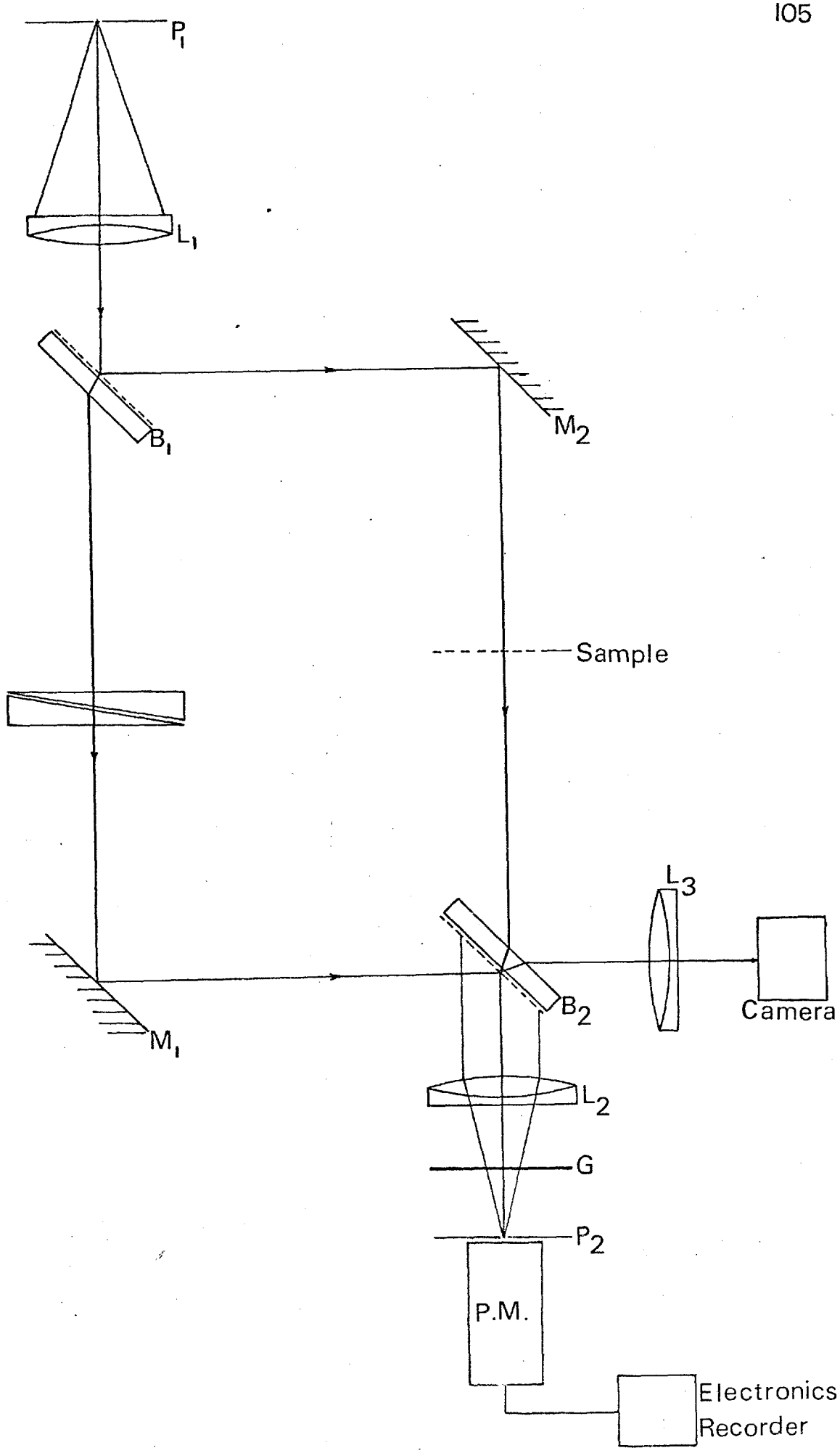


Fig 5 Interferometer and Processor Combined

CHAPTER V - THE PHOTOGRAPHIC EMULSION

The use of photographic materials as storage media has been exploited since the beginning of the century. In the last decade the laser has become an indispensable tool in many scientific laboratories and with it has appeared a new series of emulsions specially designed for holography. However for many operations performed by a laser, a variety of conventional emulsions will do. One of these operations is that of signal processing. This has been discussed in the first chapter. The object of this work is to explore some of the properties of the photographic process with particular reference to its application for optical means of processing data.

The photographic process has several advantages as well as disadvantages over other methods, such as photochromics and photoresists, of recording data. For many cases we are limited by the intensity of light falling onto our recording material; in these instances we require material with a high sensitivity. As regards speed and quantum efficiency, the photographic material is far better than photochromics and photoresists. Further factors which are relevant and should be taken into account are listed below. They cannot all be strictly regarded as advantages or disadvantages of a particular input mechanism, they should however be measured and used as a guide to the choice of recording material used.

(a) Resolution capability

(b) Dynamic range, possibility of linearity

(c) Wavelength sensitivity

(d) Good response (accurate reproduction)

(e) Permanence of recording

(f) General overall adaptability and ease of use

(g) Time of processing (real time or otherwise)

(h) Effect of light scatter

(i) Signal to noise ratio

(j) Irregularities in material and supporting base

(k) Influence of strain, development damage

(l) External effects (temperature, humidity etc.)

throughout the whole of the process.

By studying these advantages and disadvantages it may prove possible to improve the process, find new uses for the photographic process and to develop a more universal means of evaluating the capabilities of the photographic medium.

We are now familiar with the use of coherent optical processors and the way in which they can display the spatial frequency content of a recorded signal. Such techniques could be employed in inspection systems if the object has a characteristic frequency structure. Photographs could be taken of the objects under inspection and the transparency placed in the input plane of the processor. Any spectra which then appear in the frequency plane and not belonging to the characteristic frequency would be the result of some defect or unrequired structure in the object. If for example, we are looking at a smooth surface such as a sheet of glass or metal, this ideally would have no frequency content. If however it is imperfectly made - it could contain scratches, turning marks or other such imperfections - then the appearance of any spectra in the frequency plane would be an indication of

the quality of the sample. In taking the original photographs we have all the optical techniques at our disposal to magnify the effect of the defect on the recorded film. It is the examination of the diffraction plane which gives us the desired information. Small changes of the complex amplitude transmission across the aperture of the input plane of the processor can cause measurable changes on the diffraction plane. An essential function of the photographic film in optical processing is to produce controlled variations in complex transmissions. It is therefore essential to have a high degree of control both at the exposure stage and in the processing of the film. Similarly, the photographic transparency does not form a perfect system and its imperfections would influence the diffraction pattern. Phase effects which are due to optical path variations in transmission are nearly always present in all current photographic transparencies. It is the object of this work to measure these phase variations and to investigate the magnitude of their effect on the distribution of light intensity in the diffraction plane. The performance of the photographic film is determined by the combined effect of the emulsion itself and the film base. The influence of each separately, and then both combined on the efficiency of the optical processing system is considered.

The choice of the type of emulsion for the investigation was limited to those emulsions which are capable of rapid processing, as it is these emulsions which are most useful in automatic inspection techniques. The main feature of these types of emulsion is that they must be hard and durable so as to withstand high temperature processing. Four such emulsions were initially chosen for this study, two Kodak,

2496 RAR, 2498 RAR and two Ilford, RX and RE recording films.

V.1. Physical Characteristics of the Emulsions

Both the Kodak emulsions are coated on a 0.1 mm polyester base, which has the advantage of being hard and readily able to withstand high temperature processing. Unfortunately, it is birefringent and such properties are undesirable for use with plane polarised laser illumination as would be encountered in a coherent processor. Further difficulties arise when immersing the film into a matching refractive index liquid. The refractive index for light travelling with the Evector at right angles to the optic axis is 1.51, along the optic axis it is 1.64.[87] The direction of the optic axis is in the direction of the manufacture of the polyester. There is also a slight random drift in the polarisation directions. There is no antistatic coating on these films. Further mechanical properties of polyesters are given in a later section. The 2498 RAR is a high speed panchromatic emulsion. Its spectral sensitivity starts falling off at about 6200Å and has a cut off at 6500Å. This makes it relatively insensitive to exposure with the 6328Å laser line. The 2496 RAR is a medium speed panchromatic with extended red sensitivity emulsion. Its spectral response starts to fall off at about 6800Å and it is thus suited for exposure with a helium-neon laser. Both the Kodak emulsions have a base density of around 0.1.

Both the Ilford emulsions are panchromatic and coated on a 0.13mm triacetate base which is permanently grey-dyed for anti-halation purposes. Their spectral sensitivity extends to about

6400Å and their base density is just over 0.2. The emulsions are thinly coated and prehardened to enable rapid processing at elevated temperatures. The emulsion thickness was checked with a phase contrast microscope and a Talyrond 200 and found to be of the order of 3 - 4 microns. Both the emulsions are of medium speed with the RE being slightly the faster. The refractive index of the base is between 1.47 and 1.48 and of the gelatin between 1.52 and 1.53, the values being means for white light.

It has already been mentioned that the performance of the photographic film is determined by the combination of the base and the light sensitive emulsion. Their optical properties will be dealt with in the later sections, however, another important aspect, especially when processing at extreme temperatures, is their dimensional stability.[42] The magnitude of the dimensional changes for a given film depends upon the chemical composition and thickness of the base and emulsion and of course on the magnitude of the phenomena causing the changes. We are dealing with two different bases, a tri-acetate base for the Ilford emulsion and polyester base, which is a thermo-plastic material known as polyethylene terephthalate, for the Kodak emulsions; both have their relative merits. On the whole the polyester base is stronger and more durable, and has a higher dimensional stability. There are two types of dimensional changes (a) those which are temporary and reversible, and (b) those which are permanent and irreversible.

(a) Reversible Dimensional Changes

These can be caused by two different types of phenomena, firstly humidity and secondly temperature. Under normal conditions

the film is kept at a constant temperature, with the humidity being more likely to vary. For our case however, we are subjecting the film to high temperatures and hence we must be aware of the thermal effects. The following figures have been obtained from Kodak data sheet RF.10. The coefficient of thermal expansion for the polyester base is about 0.001% per degree fahrenheit; for the triacetate base it is just over 0.003% per 1°F. The humidity expansion or contraction of the film is caused by a gain or loss of moisture from the air with which it is in contact. The relative humidity of the air determines the moisture content, and thus the size for a given film. The humidity coefficient of size change per cent per 1% R.H. is 0.002% for the 0.1mm thick polyester base and 0.007% for the 0.13mm triacetate base. For preservation the films are normally stored under cool conditions such as refrigerator and hence when they are brought out for use they are subject to a change in relative humidity and temperature. It is thus advisable to take the film from its storage and leave it to reach an equilibrium with the room conditions before taking any photographs - at least one hour is required for this to occur.

(b) Irreversible Dimensional Changes

These can occur due to the processing and also ageing of the emulsion. During the processing the film swells and then contracts during the drying. If, after processing and drying it is brought back to equilibrium with the same relative humidity and temperature conditions as existed before processing, a small net change can generally be found. If after processing it is not brought back to the same conditions then of course, the change is increased.

This effect is more marked when using elevated temperatures for processing.

Permanent shrinkage during the storage before exposure does occur but it is relatively unimportant since the image has not yet been recorded. After exposure the film continues to shrink; if the film is processed within a few days this effect can be neglected unless it has been stored in extreme conditions. For the triacetate base the ageing effects arise from the gradual loss of solvents and plasticisers used in its manufacture; the polyester base does not contain solvents or plasticisers and is thus free from this effect. Another of the effects is that of plastic flow of the base due to the contraction of the emulsion; this is present in both types of base but is less marked in the polyester base. The third ageing effect is the release of mechanical stresses and strains which originate during the manufacture of the base. Both types of base would be prone to this kind of ageing, the extent is entirely dependent on the manufacture.

Storage effects have therefore some influence on the amount of shrinkage. Heat will accelerate all the three ageing effects mentioned so that the film should not be stored at elevated temperatures. Similarly, extremes of relative humidity will cause permanent changes. Low relative humidities will cause a greater contraction of the emulsion and hence increase the plastic flow, high relative humidities will increase the escape of residual solvents and plasticisers.

All the films used were stored in a refrigerator and before use they were removed and conditioned to the laboratory atmosphere for at least three hours. The film was then processed within a day and when possible all measurements were taken within two weeks.

V.2. Emulsion Characteristics

The characteristic curves were determined for the four emulsions; these can be seen in figs. 1 - 7. The illumination for these curves was provided by a tungsten iodine lamp behind a diffuser. The exposure was varied by varying the distance from the source to the negative according to the inverse square law and keeping the exposure time constant. Timing for the exposure was achieved by using an electronic shutter. Density measurements were performed using a white light densitometer, and those shown are due to the emulsion and base. Values of amplitude transmittance are for the emulsion only.

Two developers were used for each emulsion and the developing temperature was varied between 20° and 50°C. In order to avoid the effects of reticulation all the solutions used in the processing were kept at the same temperature to within $\pm$ 1°C by having the processing tanks immersed in a thermostatically controlled water bath. The developing solutions were used once and then discarded. One development consisted of the number of exposures necessary for one characteristic curve (ten) developed in 300cc of developer.

For the Kodak emulsions two developers, D19 and D76 were recommended by the manufacturers. It can be seen from the characteristic curves that D76 is much slower than D19. The development temperature for these emulsions was varied by 10°C steps from 20°C to 50°C. The development time at 20°C was that recommended by the manufacturers to give a gamma of unity (D - log E curve). The times of development for the other temperatures were adjusted so as to give as near as possible the

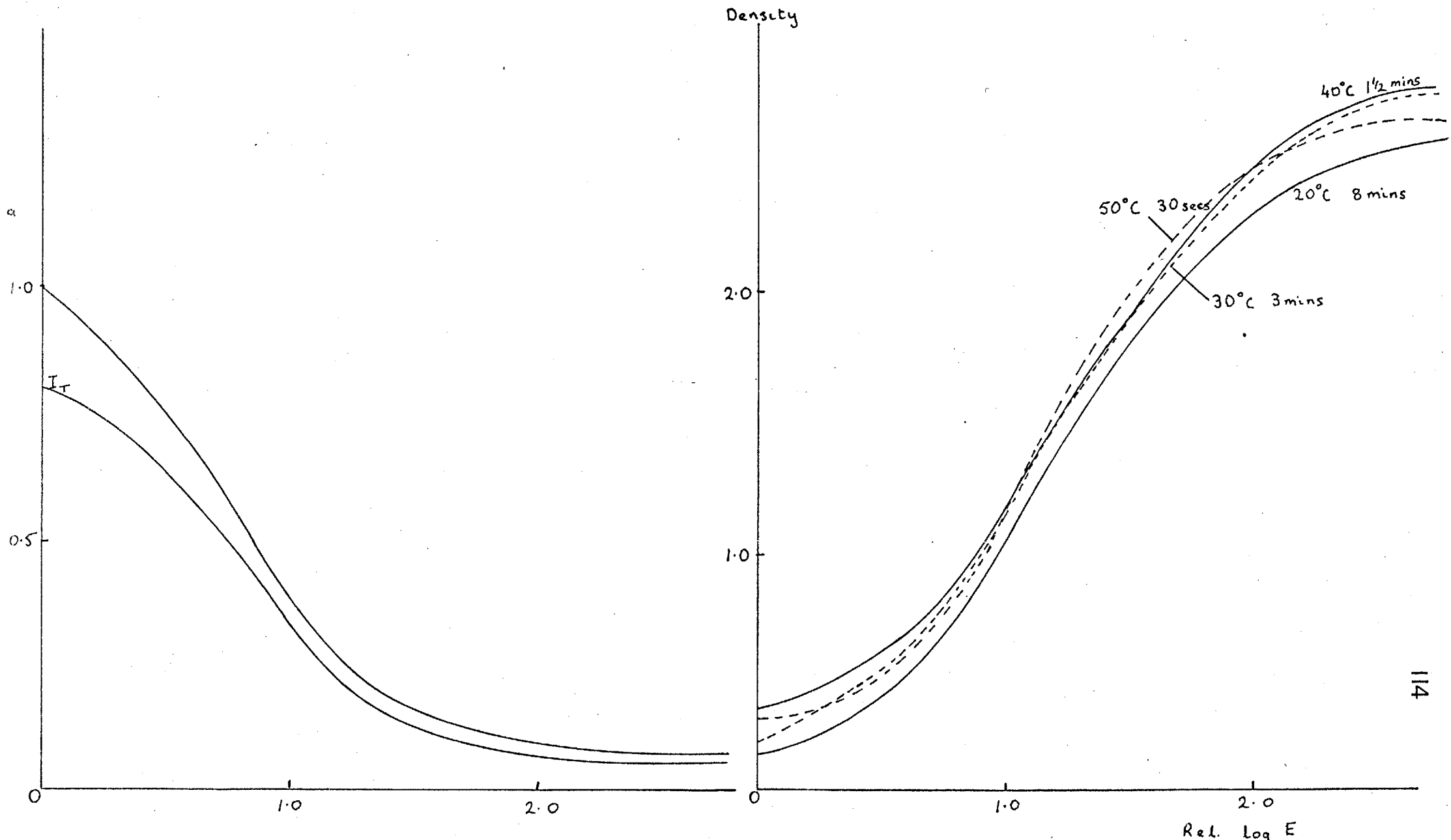


Fig 1 2496RAR D19

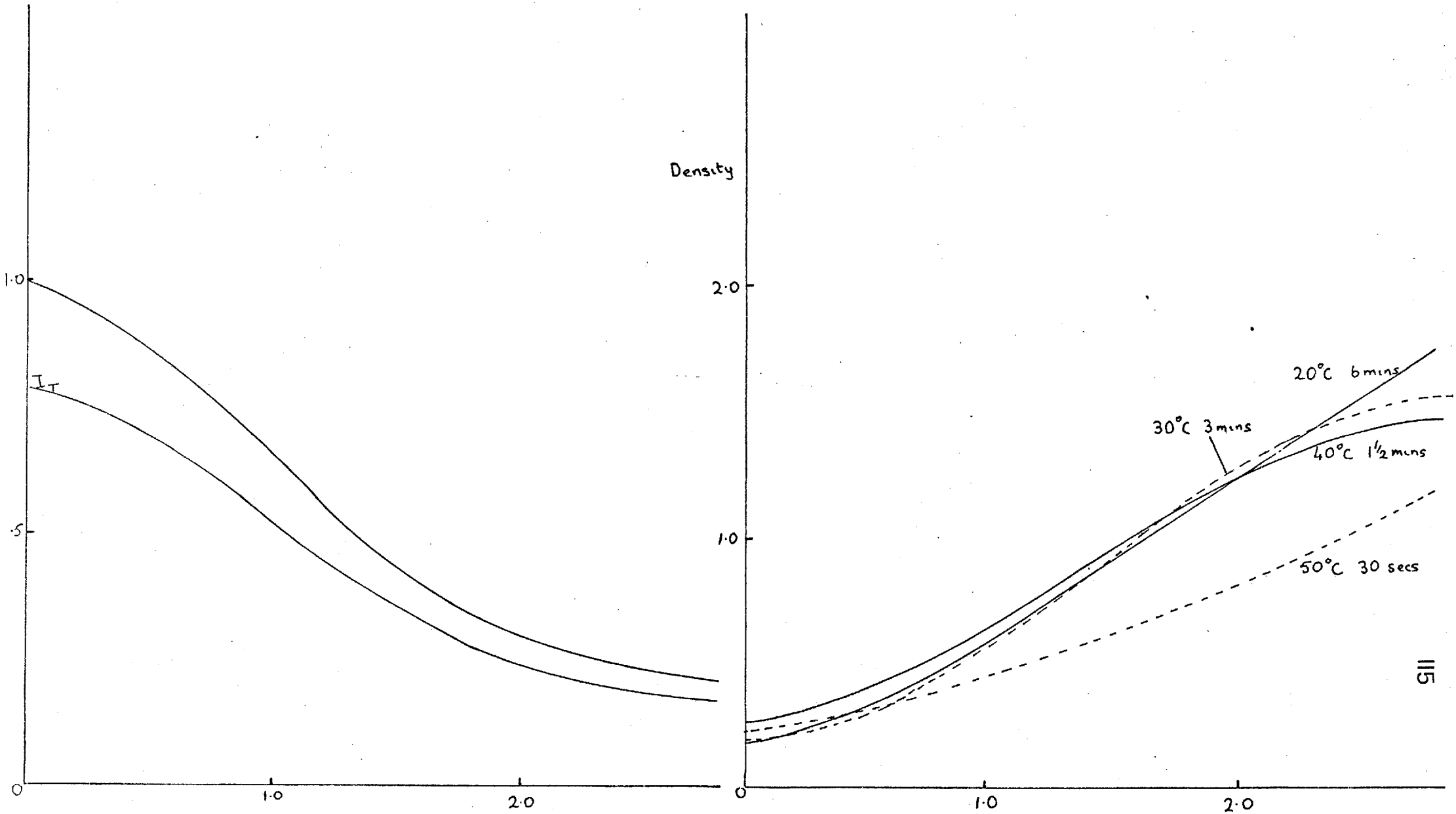


Fig 2 2496RAR D76

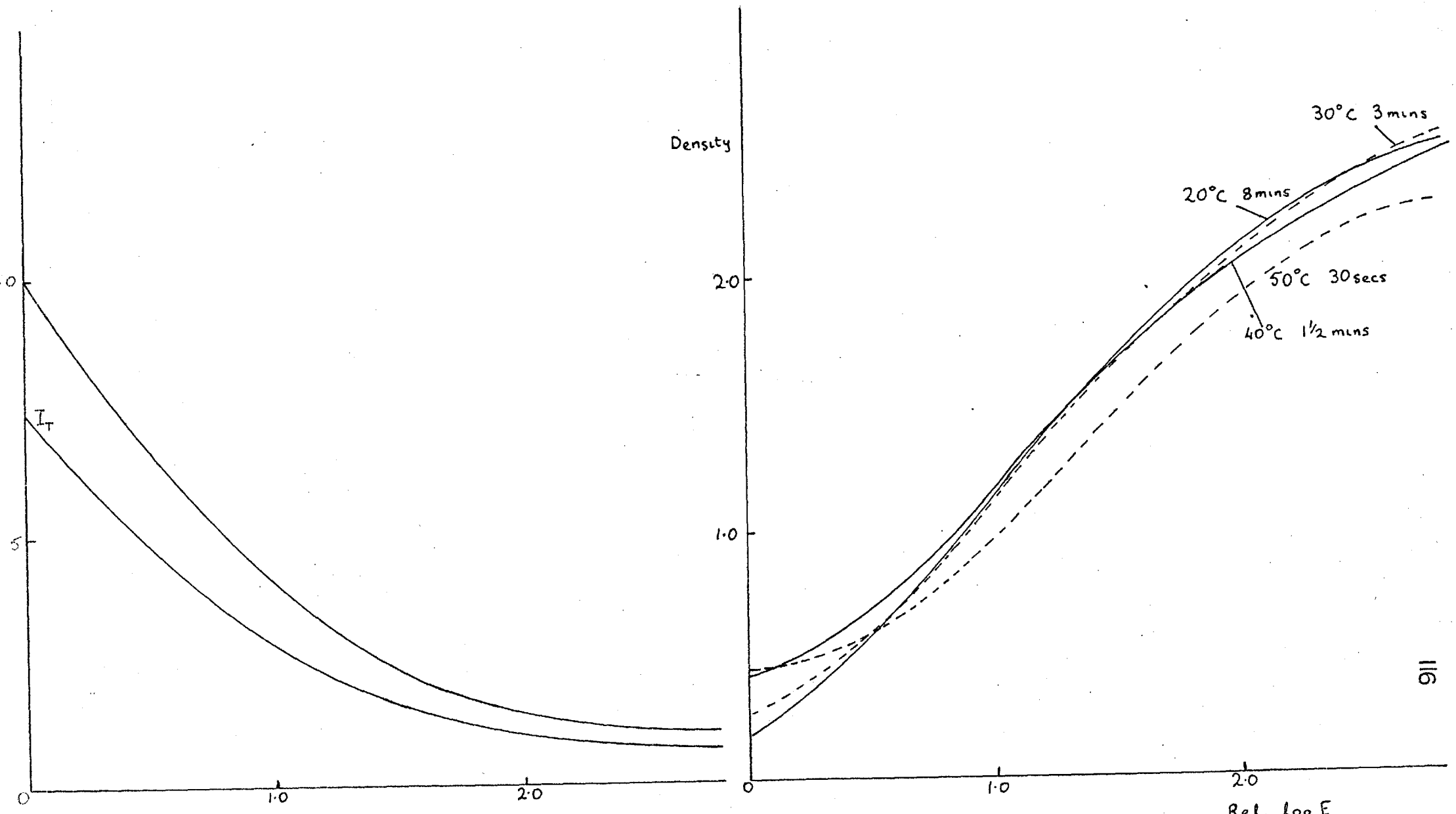


Fig 3 2498RAR D19

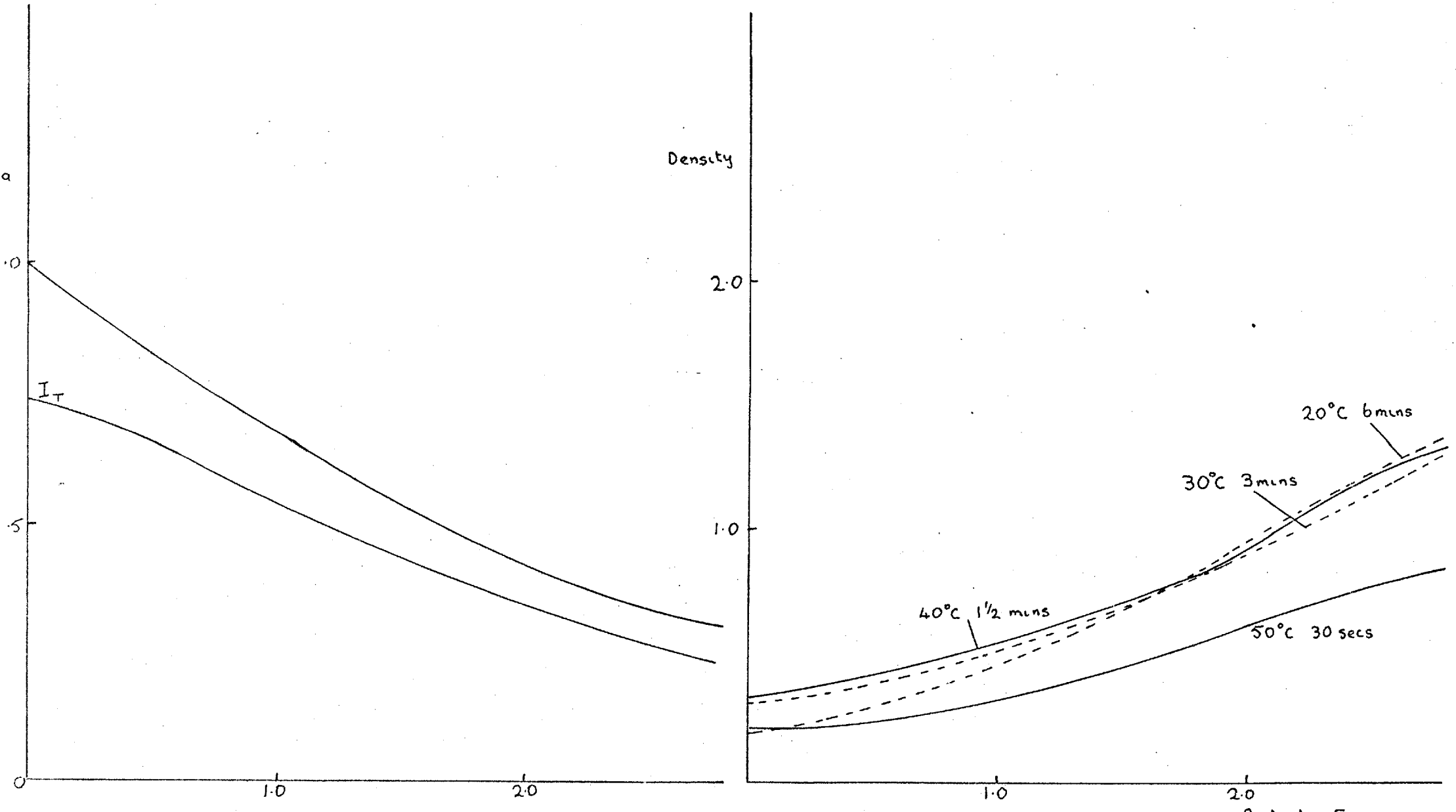


Fig 4 2498RAR D 76

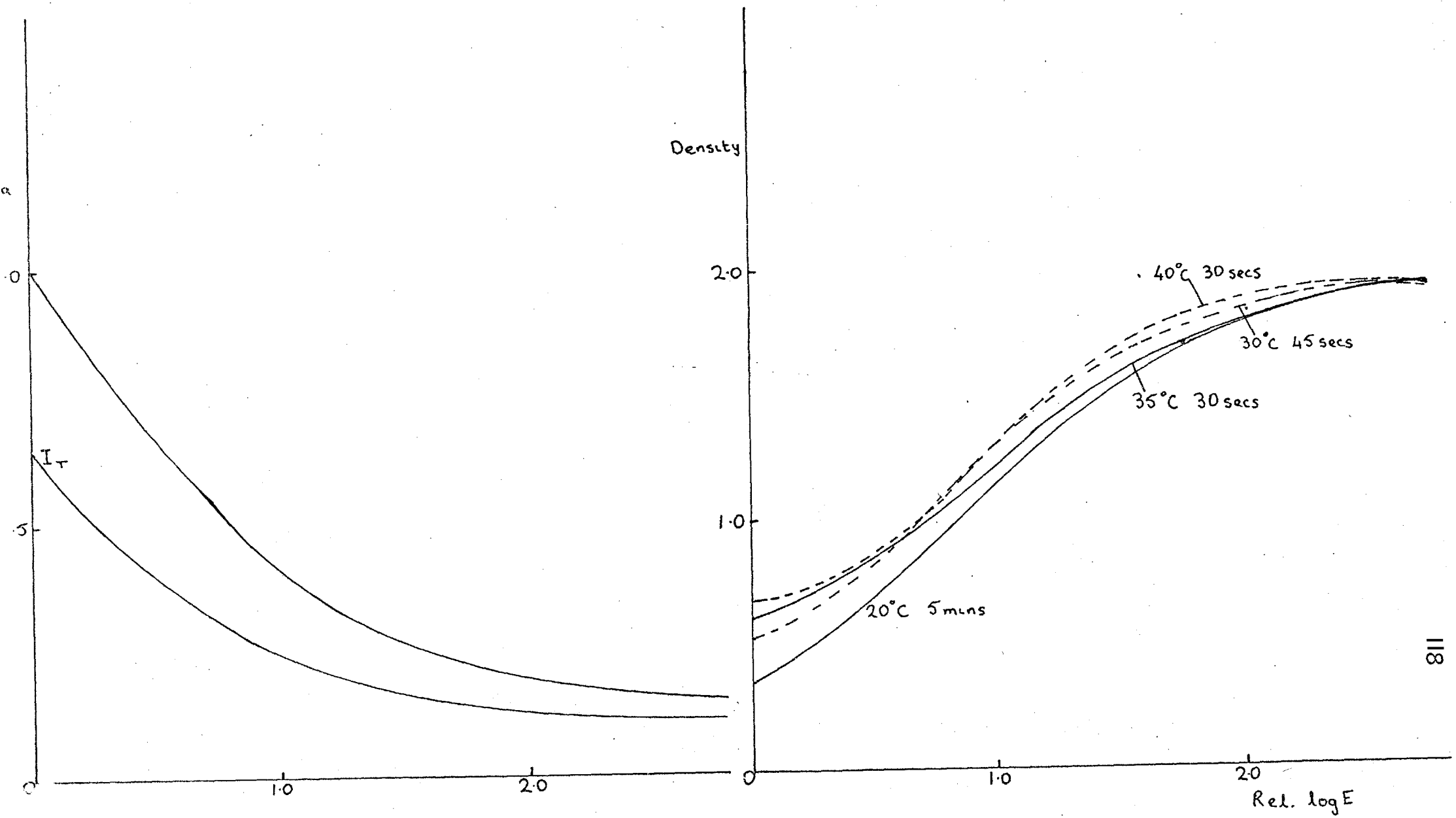


Fig 5 RX Phentrace

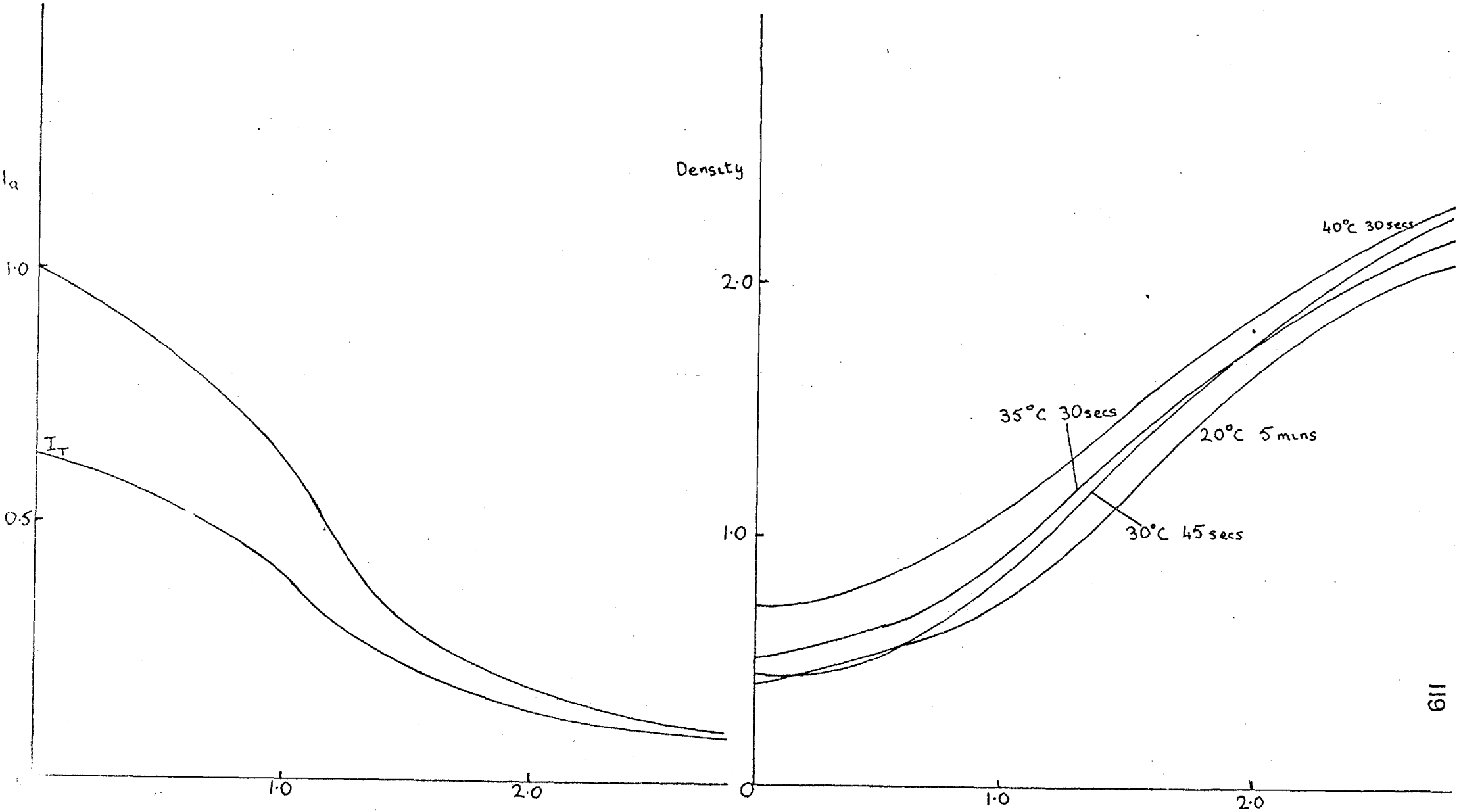


Fig 6 RE Phentrace

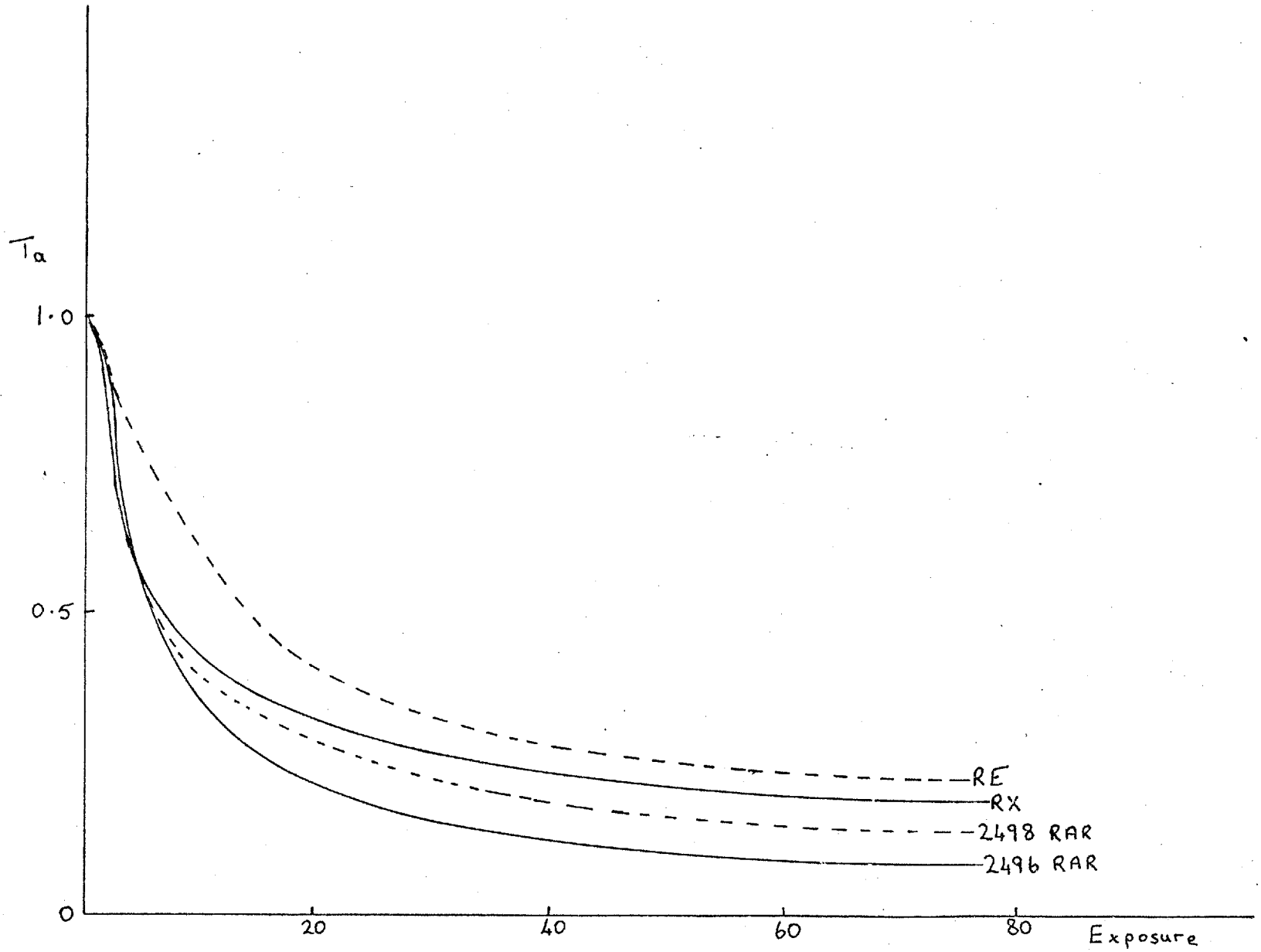


Fig 7 T_a-E Curves

same characteristic curve as at 20°C.

For the Ilford emulsion ID19 and phentrace developers were recommended. It was found that phentrace was the faster developer and as these emulsions are designed for rapid processing a fast developer is essential so ID19 was discarded. As for the Kodak emulsions the developing time was adjusted at each temperature to give a gamma of unity. The combination of the Ilford emulsions with phentrace developer required much shorter processing times than the Kodak emulsions with D19 developer. As a comparison, 30 secs developing at 35°C for the Ilford emulsions was equivalent to 30 secs at 50°C for the Kodak emulsion. When the phentrace developer was raised to 40°C it was found necessary to change the dilution from 1:4 to 1:5 and keep the developing time at 30 secs, as with the standard processing equipment available it was not practical to have a developing time less than 30 secs and be confident of the results. At the higher temperatures fogging began to occur; to reduce the effect of this 0.3gm per litre of benzotriazole was added to the developer.

The Ilford RX and RE recording films were designed in conjunction with the Kelvin Hughes rapid processing system.[18, 19] Typical processing times in this system are from 3 to 25 secs. The emulsions are passed through the processing tanks at a constant speed where a continuous stream of processing chemicals at an elevated temperature are poured onto them. Thus the decomposed chemicals produced by the reaction of the developing agents on the silver halides are never in contact with the emulsion, enabling reliable rapid processing. One of the developers sometimes used in this system is ID57; this

was tried and found to be by far the fastest developer used. The developer was used on both Ilford RX and Kodak 2496 RAR emulsions and it was attempted to obtain the same curves as for the other developers (D19 and phentrace). At a dilution of 2:1 and a temperature of 20°C the developing time for the RX film was only 6 secs and for the 2496 RAR film, 8 secs. The results were a mean of a series of developments. It must be realised that the control during the development was very inadequate firstly due to the inaccuracy in timing and secondly due to the inefficiency in agitation of the developer for such a short period. The results were repeated at a dilution of 1:2 and a temperature of 20°C, the required developing times were now 30 secs for the RX film and 35 secs for the 2496 RAR film.

The general trend from all these experiments showed that the Kodak emulsions could produce a higher saturation density and could be developed to give a higher contrast than the Ilford emulsion, but the processing time of the Ilford emulsions was shorter than for the Kodak emulsions when developed to give the same contrast.

The chemical composition of developers D19, D76 and ID19 can be found in the Kodak and Ilford data manuals but for completeness are given below. That of ID57 is not listed but its composition has been obtained and is also given below. The chemical composition of phentrace is not listed nor has it been possible to obtain the information from the manufacturers. It is known to be a highly active Phenidone-hydroquinone developer of medium-high contrast.

D19 is a high contrast developer made up as follows:-

Metol	2 gm
Sodium Sulphite (anhyd)	90 gm
Hydroquinone	8 gm
Sodium carbonate (anhyd)	45 gm
Potassium bromide	5 gm

to one litre of distilled water.

D76 is a medium contrast fine grain developer

Metol	2 gm
Sodium sulphite (anhyd)	100 gm
Hydroquinone	5 gm
Borax	2 gm

to one litre of distilled water

ID19 is a medium-high contrast developer

Metol	2.2 gm
Sodium sulphite	72 gm
Hydroquinone	8.8 gm
Sodium carbonate	48 gm
Potassium bromide	4 gm

to one litre of distilled water

ID57 is a medium contrast highly active developer

Sodium sulphite (anhyd)	118 gm
Potassium bromide	5.9 gm
Sodium hydroxide	29.5 gm
Hydroquinone	47.2 gm
Methyl phenidone	3.2 gm

to one litre of distilled water.

V.3 Phase Measurements

When dealing with photographic emulsions in a coherent optical processor we have already indicated that we are interested in the complex amplitude transmission of the emulsion. Photographic film consists of light sensitive silver halide grains held in a gelatin glue and supported on some form of base. In this section we will be looking into the phase variations present in two of the rapid access emulsions. One emulsion was chosen from each manufacturer, these being the Kodak 2496 RAR and Ilford RX emulsions. The phase variations were recorded using the Mach-Zehnder interferometer.

Initially two types of interferograms were taken. One with the interferometer set up for fluffed out fringes; this gave a quick indication of the shape of the emulsion. The other arrangement with about ten tilt fringes introduced over the 35mm frame, gave a more detailed picture of the emulsion performance. Photographs of the interferograms were taken on 35mm film using a 135mm telephoto lens.

Before examining the phase variations of the silver emulsion the phase effects of the supporting base were determined. The 2496 and RX emulsions were removed from their polyester and triacetate bases by immersing them into a weak solution of caustic soda for about 2 mins. The bases were then placed into the interferometer and the interferograms taken. The interferograms showed that the statistical distribution of phase variations in the two kinds of base were similar. The fluffed out interferograms showed that only small portions, a maximum of about 1 sq.cm area, of the base

were flat to a quarter of a wavelength.

Analysis of the tilt fringe interferograms was achieved by comparing the fringes obtained through the base alone with an ideal set of test fringes (straight line fringes) obtained without the base. In many cases the number of fringes on the two sets of interferograms were not the same. This is because the base was slightly wedged or contained some kind of bulge or indentation. Readings of the amount by which these two sets of fringes differed were taken at equally spaced intervals along each test fringe, the readings being taken to 1/10 of a fringe. The wavefront deviation was found and the standard deviation of these wavefront deviations was calculated for the whole frame. Over one hundred interferograms were taken for each base and a histogram of the standard deviation was plotted. The observations were repeated, this time using developed but unexposed emulsions. These interferograms showed the phase variations produced by the base and gelatin. The results were analysed as before and the histogram of the standard deviation plotted.

Figs. (8) and (10) give the comparison of the two types of base. The results show us the amount of wavefront deformation we can expect using an aperture of 10sq.mm in area. The optical path variation is distributed over a very wide range; in both cases the mean of the variations is 0.24 λ for the polyester base and 0.26 λ for the triacetate base. The significance of such a small difference is not very great; it can quite easily be attributed to the dimensional stability of the two types of base as discussed in section V.I. Using this reasoning it could be expected that the polyester base would be slightly better

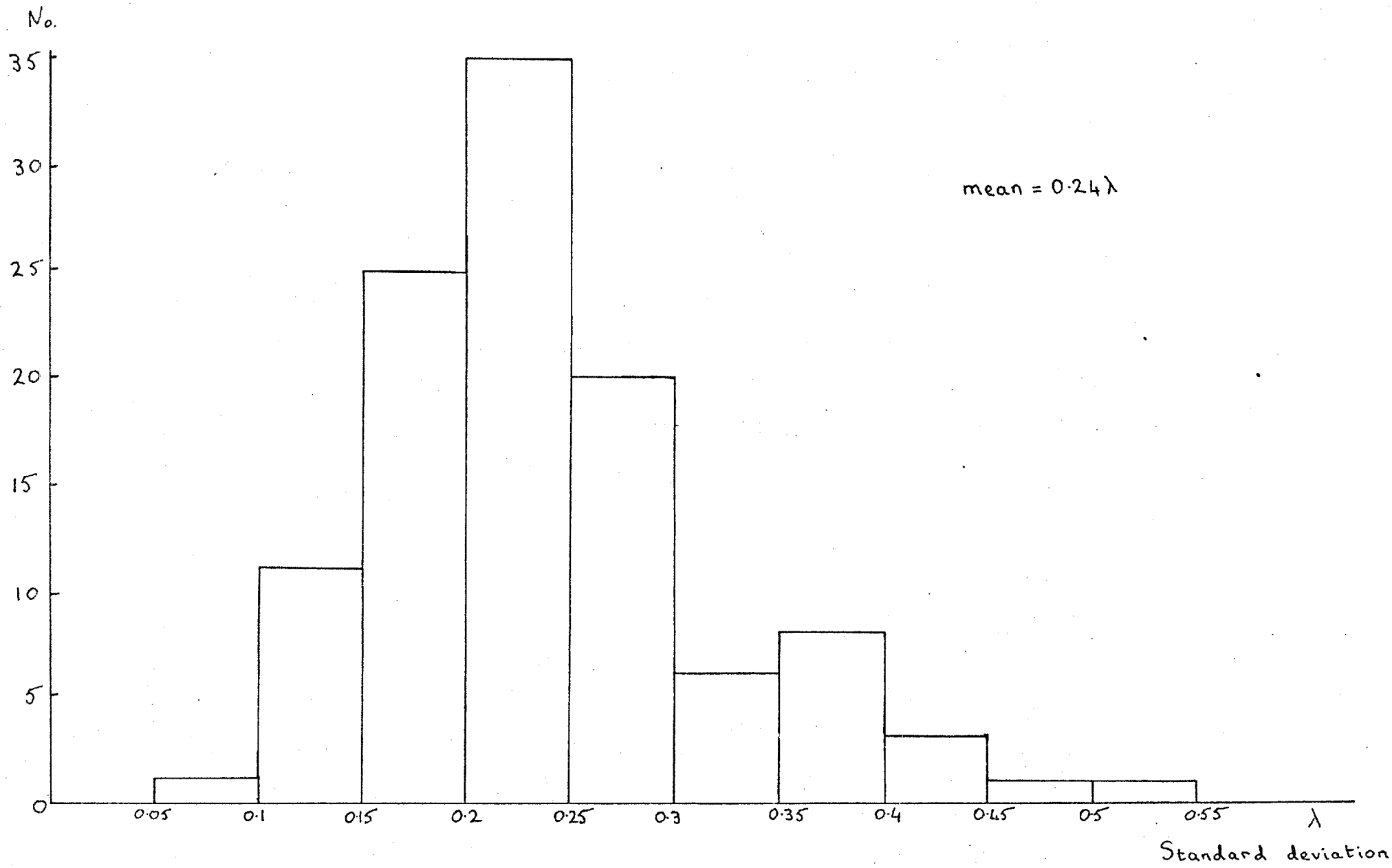


Fig 8 2496 RAR Base

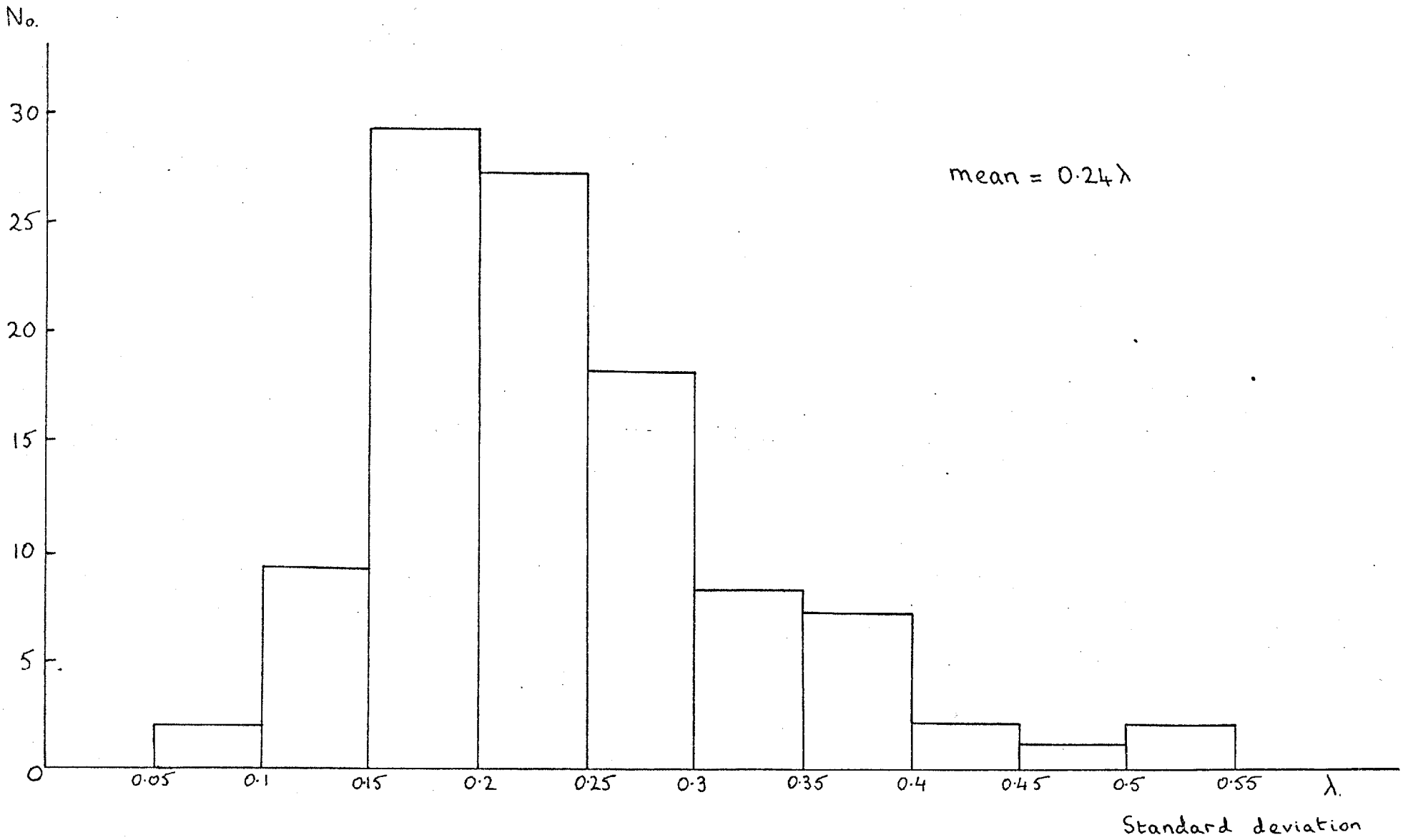


Fig 9 2496RAR Base + Gelatin

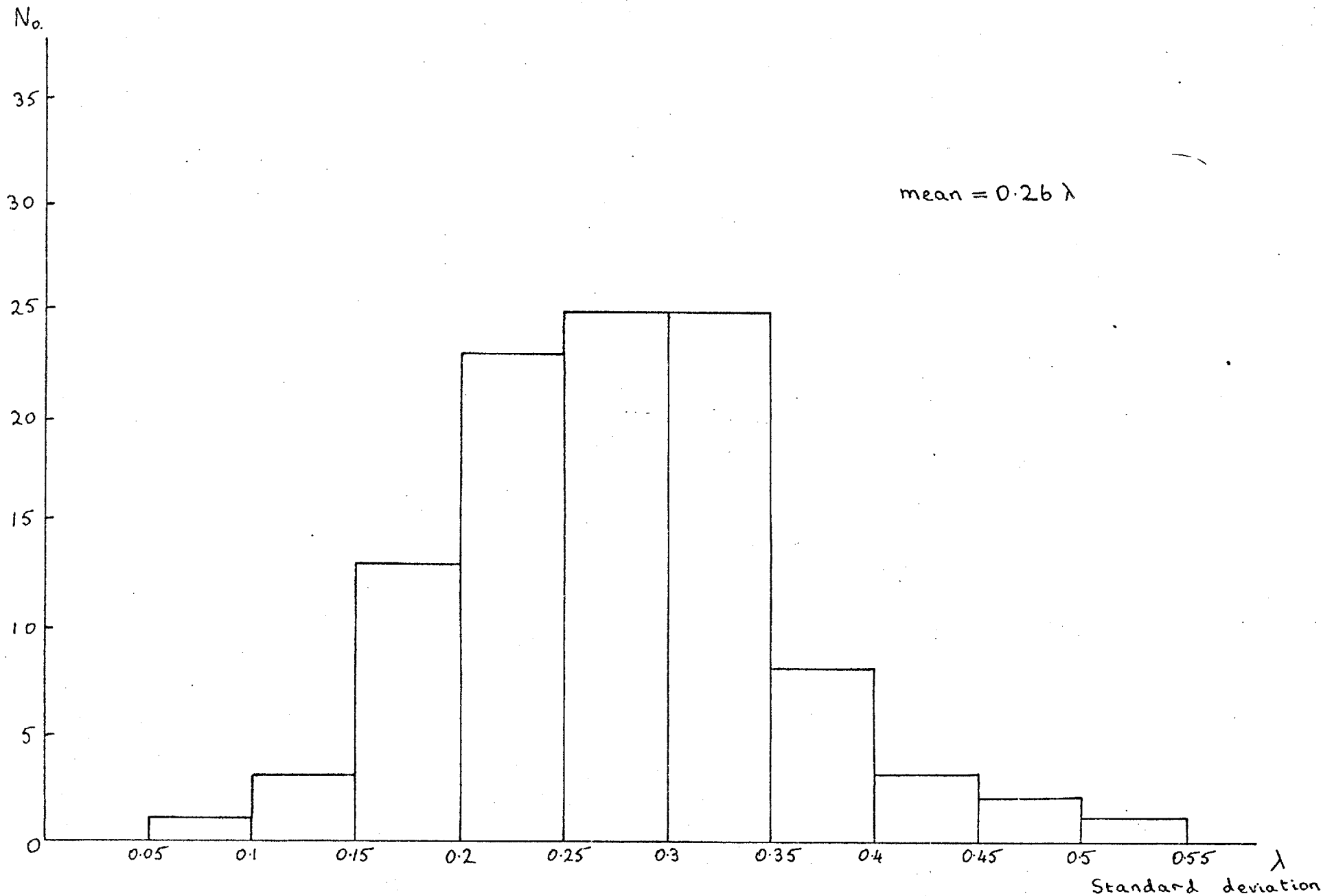


Fig 10 RX Base

especially when one considers the ageing effects due to the necessity of obtaining the film in bulk and thus having a relatively long storage time. It can therefore be concluded that the distribution of phase variations in the two types of base are very similar especially when kept under proper conditions. We are now in a position to impose some kind of tolerance criterion on the quality of the base. For the present case we will use the Rayleigh criterion which states that no more than one quarter of the wavelength of light of optical path difference is permissible over the aperture. We have in effect been looking at the optical path differences over a 10sq.mm aperture and from the results we see that about 50% of such areas are within this limit.

From a comparison of figs. (8) and (9) it can be seen that the gelatin has no adverse effect on the optical path variations. The two histograms are very much the same and the mean of the optical path variations is 0.24 λ in both cases.

The next stage was to examine the effects of uniformly exposed densities and developing temperatures. From the information obtained from the manufacturers it could be expected that the emulsions were similar. This seems to be a reasonably good assumption since both emulsions have been designed with the same objective in mind. Therefore, rather than duplicating a whole series of lengthy experiments, it was decided to concentrate all the work on one emulsion. Since the developing times of the RX emulsion with its recommended developer were slightly shorter than the 2496 RAR emulsion, the latter emulsion was discarded.

Strips of Ilford RX emulsion were exposed to various uniform

densities from just above fog level up to saturation density of two. The extremes of developer temperature were used, from 20°C up to 50°C. These strips of emulsion were then in turn placed into the interferometer and photographs of their phase variations, using tilt fringes, were taken. Fluffed out interferograms were also taken for some of the strips. The emulsion was then completely removed from the base and the same set of photographs were taken with the base alone. As before these fringes were compared with a series of test fringes, taking equally spaced readings, measured to 1/10 of a fringe, along each fringe. Fig. 11 shows the distribution of phase variations for linearly recorded uniformly exposed densities in range 0.5 to 1.5. Half of the exposures were developed at 20°C and the other half at 50°C and the histogram is the sum of the two. The mean of these results gives us a phase variation of 0.13 λ which is 0.05 λ more than that of the base alone. This is significantly different but is only a small portion of the total variation.[38] If we were to split this histogram according to the developing temperature we should get similar results to those of figs. (8) and (10). The mean of the emulsion developed at the lower temperature is 0.30 λ and at the higher temperature 0.32 λ . In this case the small difference could well be attributed to reticulation effects, although the developer stop bath, fixer and first wash temperatures were all the same, the air temperature was 30°C below this temperature and the emulsion was still subject to large thermal changes and hence subject to expansion and contraction.

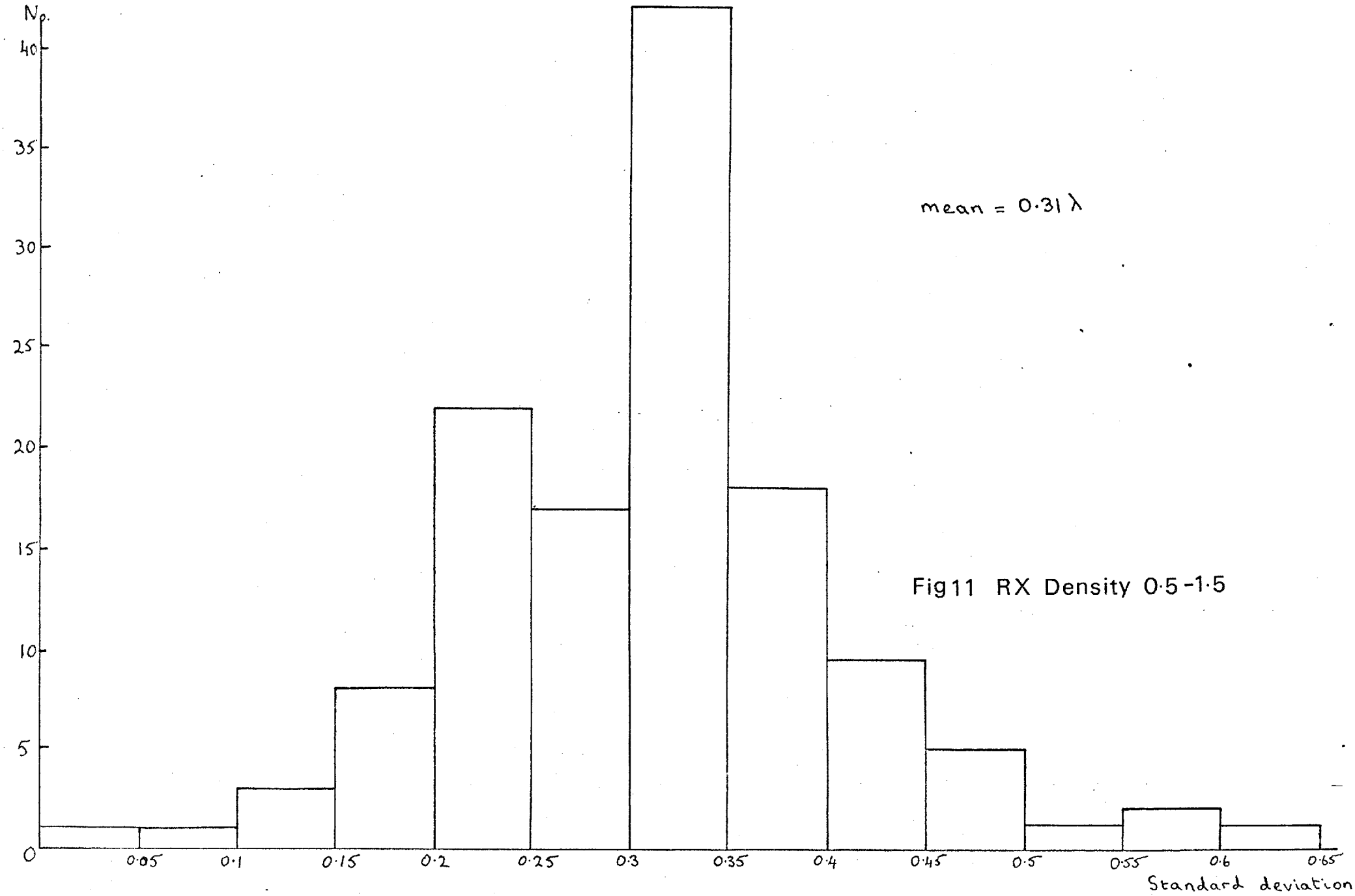


Fig 11 RX Density 0·5 -1·5

Fig. (12) shows the distribution of phase variations for film densities of two. The mean of these results is again 0.31 λ ; it is however of some significance that relatively few portions of these emulsions satisfied the Rayleigh tolerance.

In order to do a complete analyis of the interferograms a computer program was written (see appendix) and a tolerance criterion equivalent to the Strehl Intensity Ratio was used. This criterion is equivalent to the Rayleigh limit but is arrived at by considering the amount of aberration required to change the intensity of the central peak at the diffraction focus by a certain amount. It can be shown that the Strehl ratio [8] is given by

$$I = \left(1 - \frac{2\pi^2 E}{\lambda^2}\right)^2 \qquad (5.1.)$$

where I = Intensity at centre of diffraction pattern with aberration / Intensity at centre of diffraction pattern without aberration

E is the variance of the wavefront aberration and λ the wavelength of light used. The Strehl tolerance is that $I \geq 0.8$. If we therefore let E be the variance of the wavefront shape obtained from the interferogram then,

$$\left(1 - \frac{2\pi^2 E}{\lambda^2}\right)^2 \geq 0.8 \qquad (5.2.)$$

giving $$E \leq 0.005\ \lambda^2 \qquad (5.3.)$$

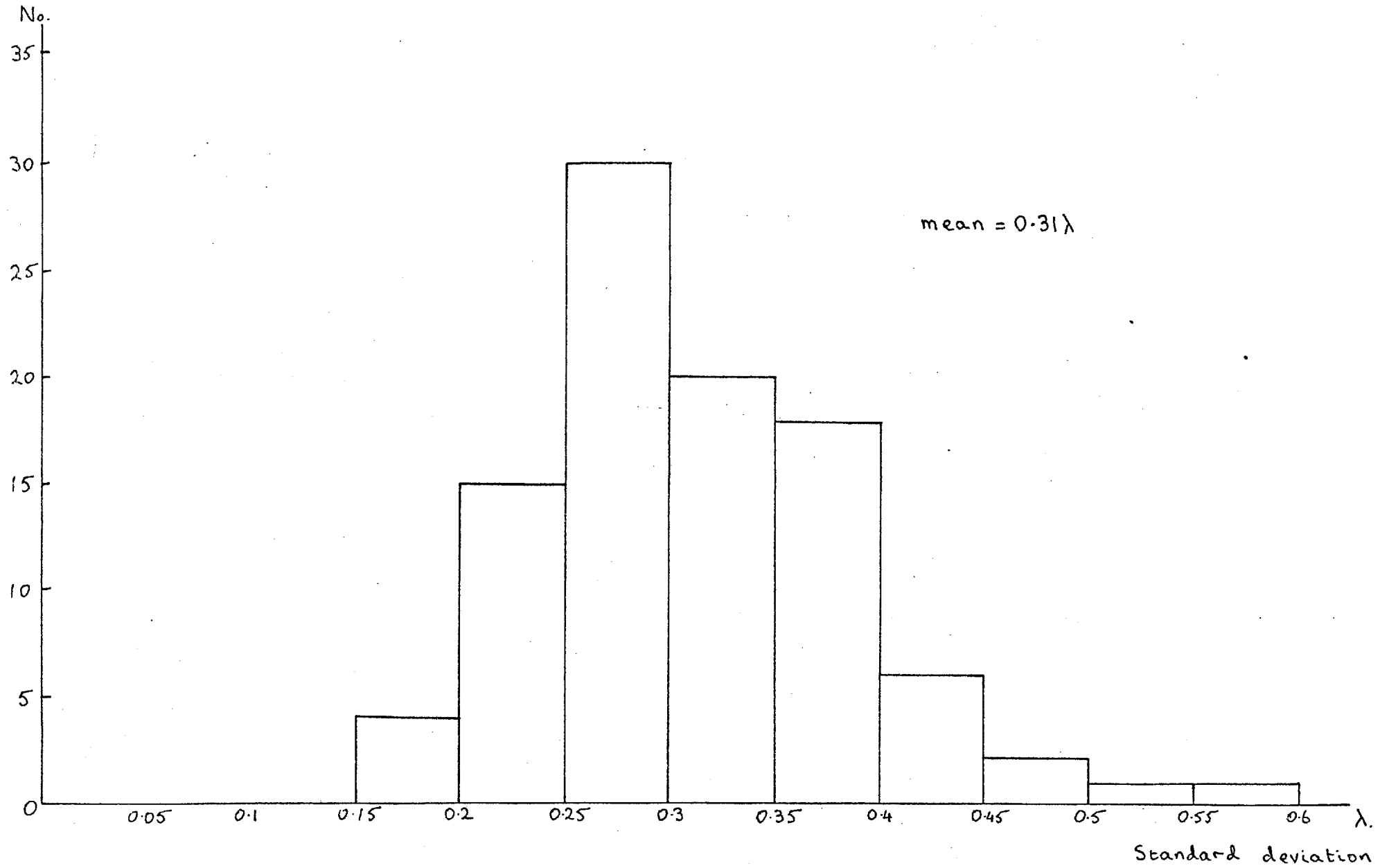


Fig 12 RX Density of Two

This criterion was used as a basis of determining good or acceptable portions of the emulsion. The computer program evaluated the variance for areas of 25sq.mm up to the full frame and the percentage of those areas within this tolerance limit for each frame; it also plotted the variance as a function of position on the emulsion. The areas were approximately rectangular in shape, resembling the 36 x 24 mm frame.

As regards the plot of variance as a function of position it was immediately seen that there was no preferred position for good or bad areas. Those parts of the film close to the perforations (within 1mm) did show rapidly varying phase changes; these however do not form part of the 36 x 24 mm frame.

From the results of the plot of the percentage of areas inside the tolerance criterion as a function of area for the different exposure and development conditions, (Fig. 13), it is seen that all the curves are of the same form, and are of a negative exponential in nature. The best fit curve was evaluated for these results and found to be

$$x = 33\ e^{-0.0098\ y}$$

where x is the percentage of areas within the tolerance and y the area of sample. This curve is a good fit for areas greater than 50sq.mm but is well below the experimental results for smaller areas. The experimental results show that as the area decreases from 50sq.mm then the percentage of good areas increases rapidly. It is seen that unless areas less than 25sq.mm are used it cannot be hoped that more than 30% will be within the tolerance limit.

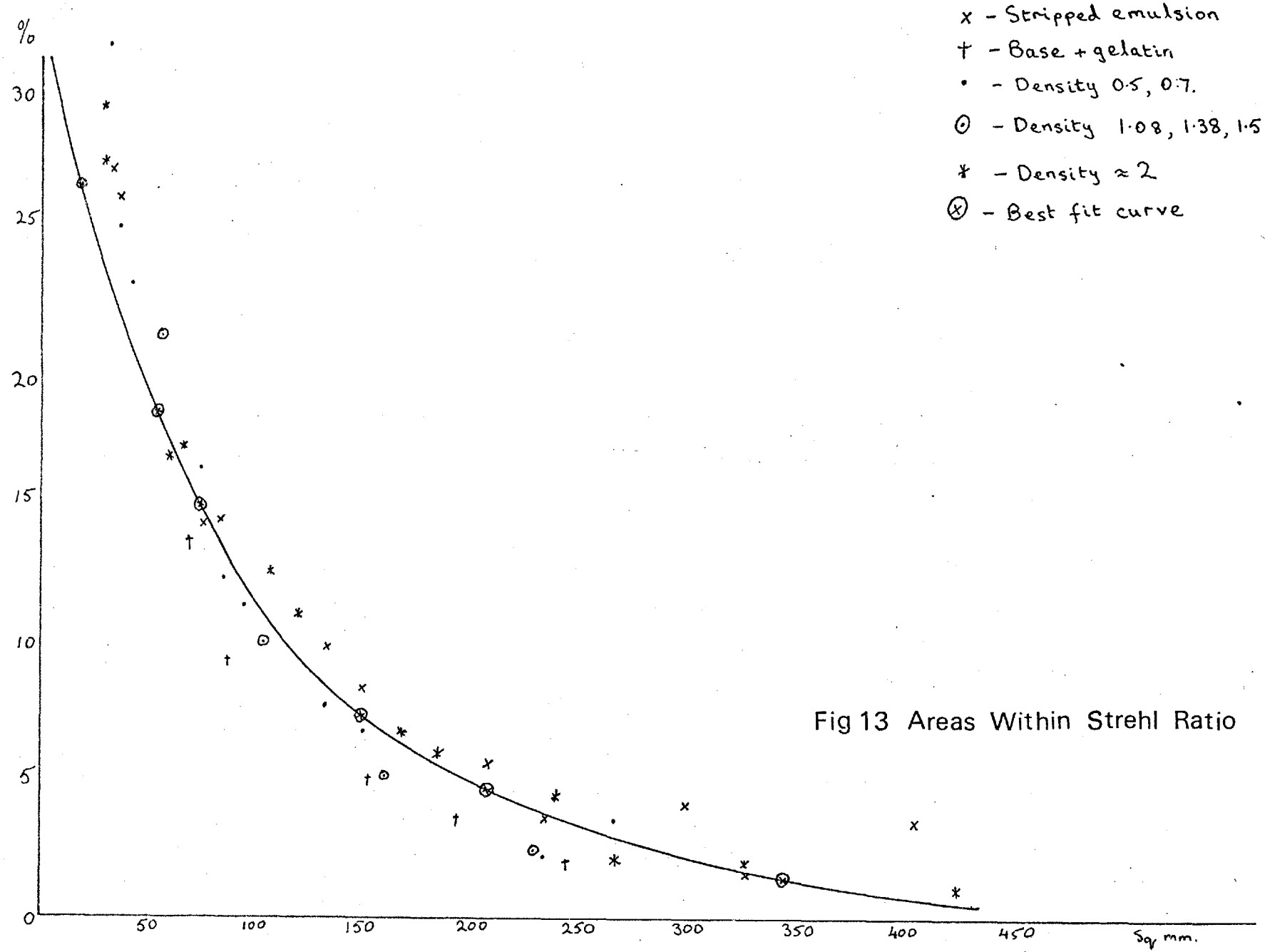


Fig 13 Areas Within Strehl Ratio

For a 50% probability it is expected that areas less than 10sq.mm need to be used. Even so, areas in different parts of the emulsion which are themselves within the tolerance limit could well have different optical path lengths in transmission.

In conclusion therefore it was found that most of the phase variation present in the film was due to the supporting base. As regards different densities it was found to a good approximation that the results were independent of uniformly exposed density. From a comparison of the fluffed out interferograms of the stripped and coated emulsions it was seen that they were similar in nature. In order to have a probability of more than 50% of choosing a good portion, areas less than 10sq.mm should be used.

V.4 Relief Images

In the previous section we examined the phase effects present in uniformly exposed density and found them to be of a random nature. In this section we will be dealing with spatially varying density, the optical path differences introduced being of a definite pattern. These phase variations are known as relief heights or relief images and have already been mentioned in II.7. The cause of these relief images is primarily the tanning action of the gelatin, and thus an important factor in the magnitude of effect is the type of developer used. The developer used for the present investigation was Phentrace which is Phenidone based and not regarded as a strong tanning developer. A further enhancing of the relief image is caused by the greater quantity of developed silver in the higher density areas, together with perhaps a refractive index change. For the

present case we are not so much interested in the cause of the relief images but their magnitude and their effect on the diffraction image. A quick check was however made to see the magnitude of any refractive index change by immersing the film into a liquid of known refractive index. From this check it was found that any change of refractive index must be less than 0.005, which would contribute only about 10% to the relief height measured. These results would agree with those of Lamberts.(50)

The Ilford RX emulsion was exposed to a spatially varying sinusoidal signal of the form

$$E(x) = Eo \cdot (1 + M \cos \omega x)$$

obtained from a two beam interference pattern. Eo is the bias exposure and M the modulation: the frequency ω of the fringes is determined by the angle between the two interfering beams. The range of exposures used was mainly on the linear portion of the D - log E curve which of course means that the density - exposure characteristics are not linear. The resulting density variation is therefore not linear, but this is of little consequence for the present study. Several sets of fringes were thus exposed onto the emulsion; these varied in frequency from 0.47c/mm up to nearly 30c/mm and peak to peak density variations up to a value of 0.95. The emulsion was then developed in Phentrace developer for 5 mins. at 20°C. The density variations were measured using a Joyce Loebl microdensitometer which had been previously calibrated. The width of the scanning slit was adjusted to be small compared to the frequency of the fringes.

The optical path variations of the relief images were measured as phase retardations using the Mach-Zehnder interferometer. The phase retardation is (n - 1) times the relief height, n being the refractive index of emulsion. The fringes were placed in one arm of the interferometer and interferograms of tilt fringes across the exposed fringes were taken in the same way as in the previous section. In order to look at the higher frequency fringes a microscope system was set up around the fringes to magnify them. These photographs showed the phase retardation as a function of density change and spatial frequency.

The results were plotted to show the phase retardation as a function of density change ΔD, where ΔD is given by Dmax - Dmin. The value of the phase retardation is measured in fringes, one fringe corresponding to an optical path difference of one wavelength, this being 6328Å. Fig. 14 shows the linear relationship between the phase retardation and the density change for the particular range of densities and spatial frequency used. Fig. 15 shows the relationship between the phase retardation and spatial frequency, for a unit density change. Here again we find that the phase retardation increases linearly over the frequency range 1 to 26.6c/mm. The results obtained seem to be in agreement with those obtained by Smith[72] and more recently, Lamberts.[50]

The emulsions were then bleached using the Kodak R-10 bleach as described by Altman.[3] This bleach consists of two solutions A and B mixed in equal parts to ten parts of distilled water.

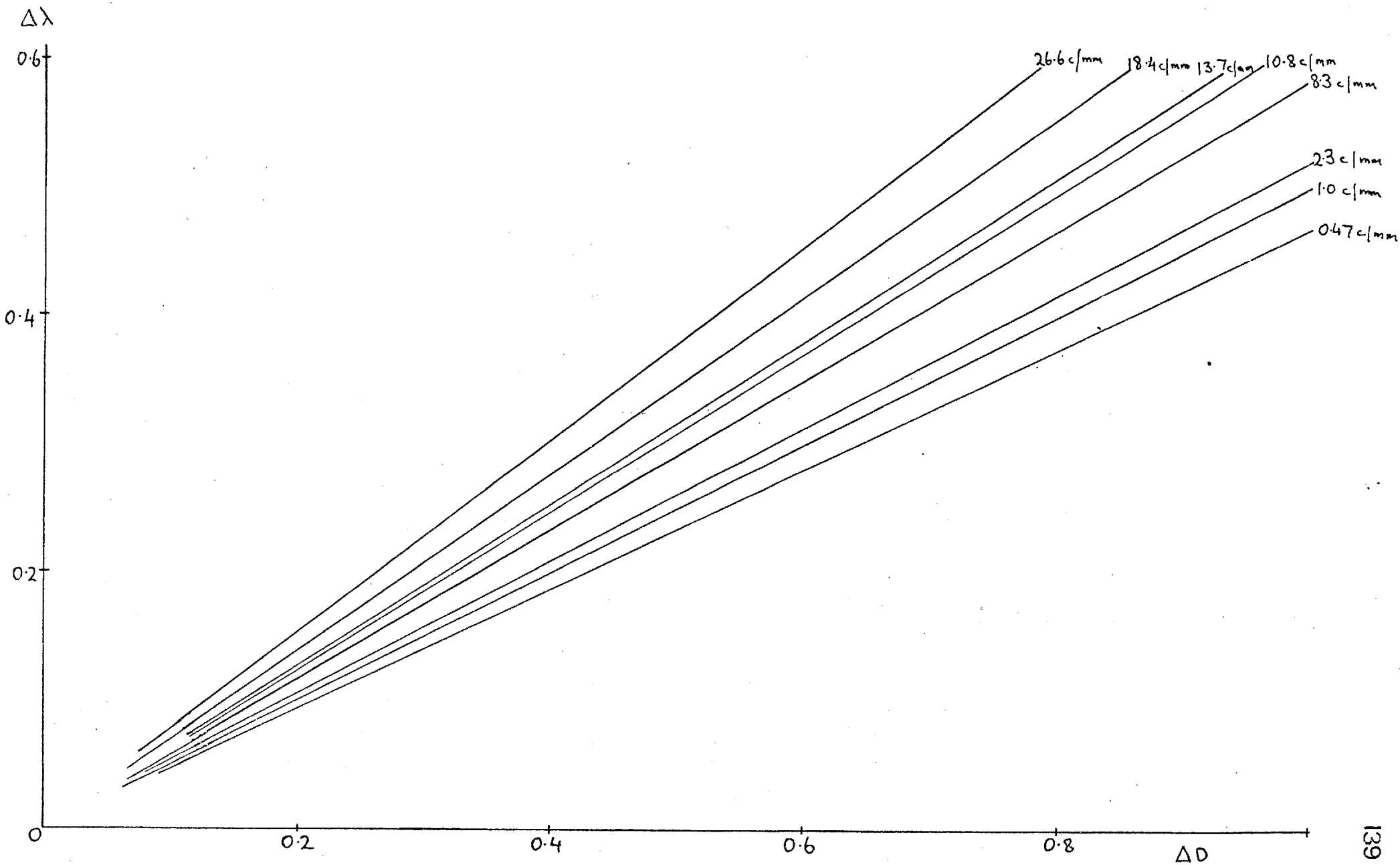


Fig 14 RX Relief Image

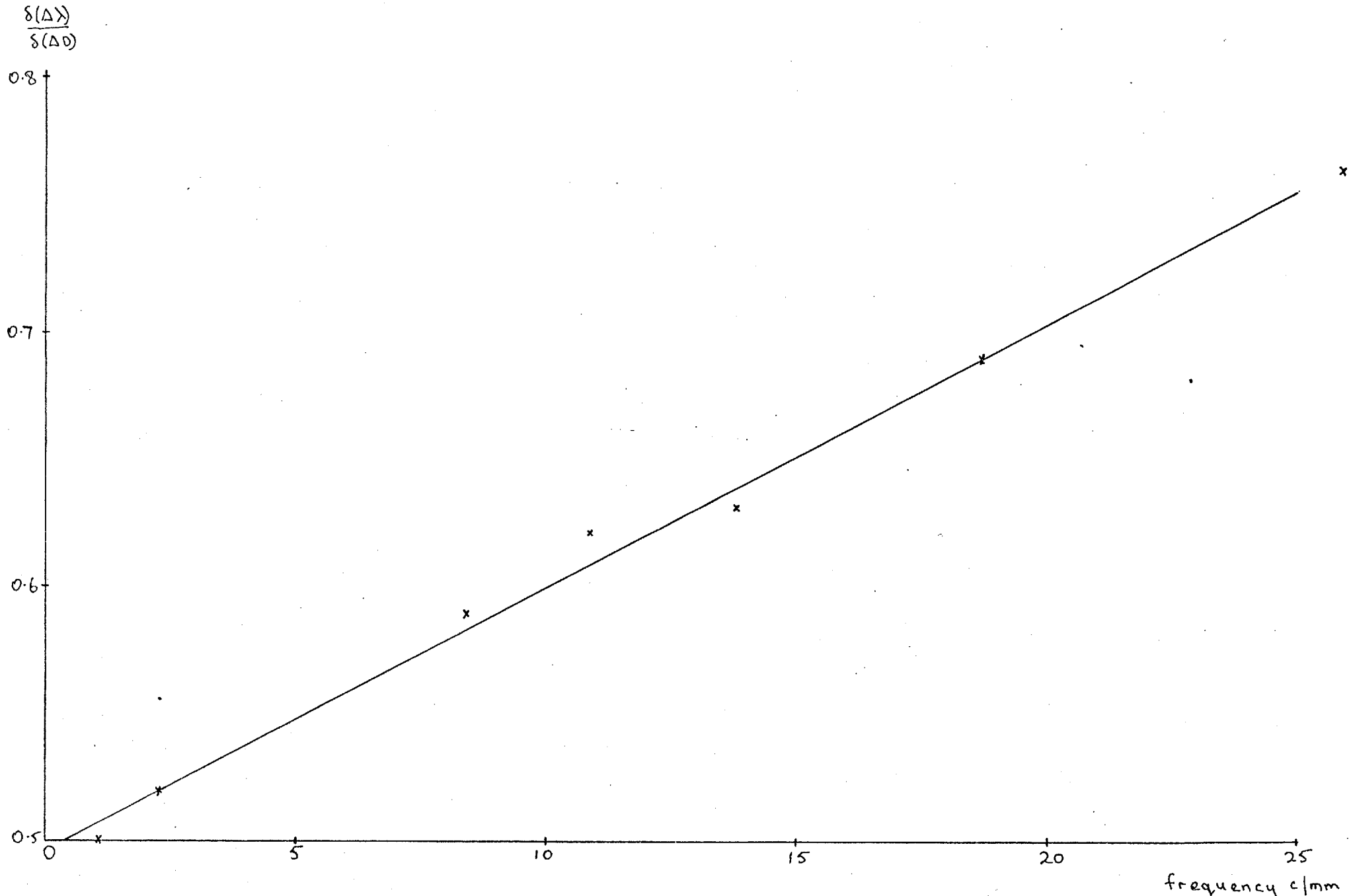


Fig 15 RX

Solution A.	Distilled water	500 ml.
	Ammonium Bichromate	20 gm.
	Conc. sulphuric acid	14 ml.
	Distilled water to one litre.	
Solution B.	Sodium Chloride	45 gm.
	Distilled water to one litre.	

The emulsions were left to soak in distilled water for about half an hour then transferred to the bleaching solution at 20°C. It could be seen that the silver had been bleached away after 2 mins. The emulsion was then washed for a further 30 mins. and left to dry slowly at room temperature.

The bleached emulsions were then placed in one arm of the interferometer and the phase retardations were measured as before. Graphs were plotted of the phase retardations as a function of original density for the various spatial frequencies, (Fig. 16). As for the silver image it was found that the phase retardation increased linearly with original density. The magnitude of the phase retardation was however only about half of that produced by the silver image. Fig. 17 shows the phase retardation as a function of spatial frequency, per unit original density change. Here it is seen that above 2c/mm the phase retardation remains constant with frequency. This result seems to be a little surprising since one would expect that the frequency at which the maximum relief height occurs would be the same for the silver image and bleached image.[72] As this frequency has not been reached for the silver image it is a little difficult to understand why it has apparently been reached for the bleached image. It could well be that this emulsion does not easily lend itself to bleaching. It is thin and

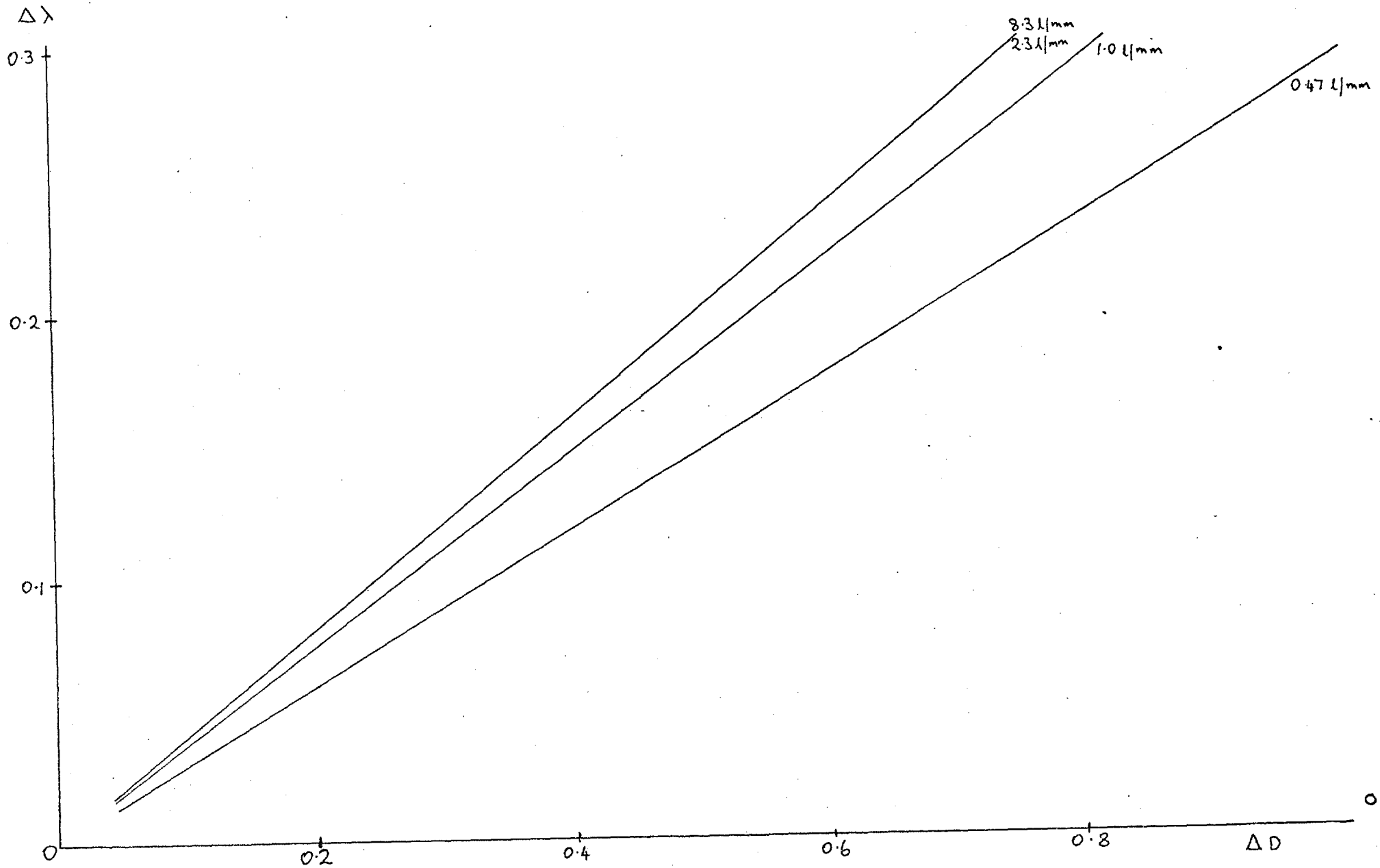


Fig 16 RX Bleached

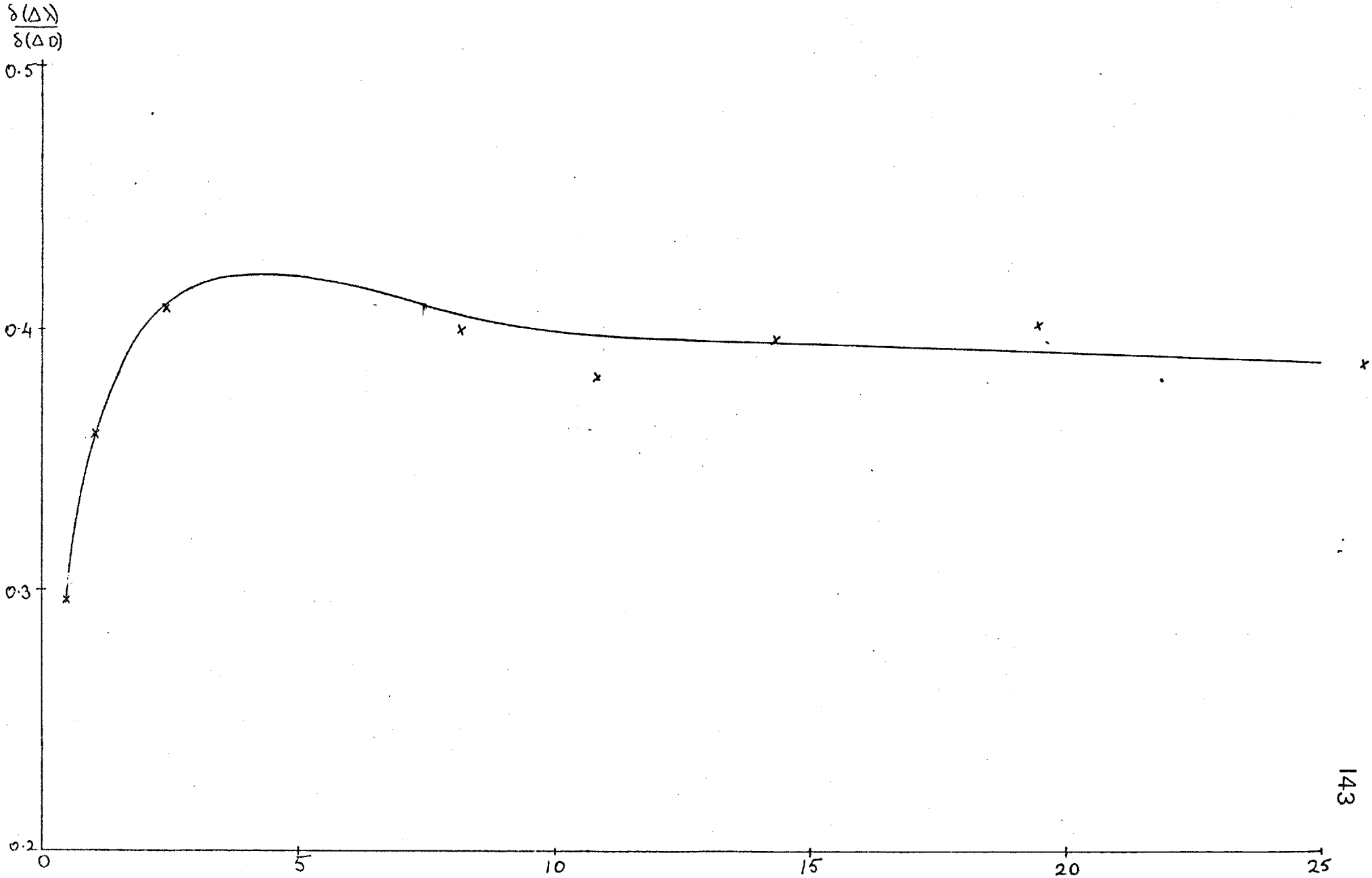


Fig 17 R X Bleached

prehardened so as to withstand hard treatment. The bleaching itself was performed by having a long time lapse between initial processing and bleaching. It could therefore be possible that the bleaching process was rather inefficient which would tend to give a rather shallow curve with no real peak. As for the present study we are more interested in the effect of the optical path variations on the diffraction plane and not in bleaching,it was decided not to pursue the matter any further, but to try to evaluate the effect of phase retardations on sinusoidal amplitude varying signals.

In II.7 we derived a relationship for the amplitudes of the zero and first order spectra in the presence of phase variations with amplitude variations. We have now verified experimentally the assumption then made that the phase retardation is proportional to the density and we are now in a position to apply eqns. 2.36, 2.37 for our present case. We have already shown, (Figs. 6, 7) chapter II how the amplitudes of the zero and first order spectra are effected by the relief image when the bias amplitude transmittance remains constant and the amplitude of the spatially varying signal increases. In normal use the film is used to record various spatial frequency signals at various bias amplitudes. We therefore wish to investigate the effect on the diffraction plane of (a) changing the bias amplitude transmittance, and (b) changing the spatial frequency, which in effect means changing the phase retardation constant, k. Figs. 18, 19 and 20 give a plot of the zero and first order spectra amplitudes for the two conditions. In these graphs the factor e^{iz} (eqn. 2.34) has been omitted. This is because it does not change the amplitude of the spectra,

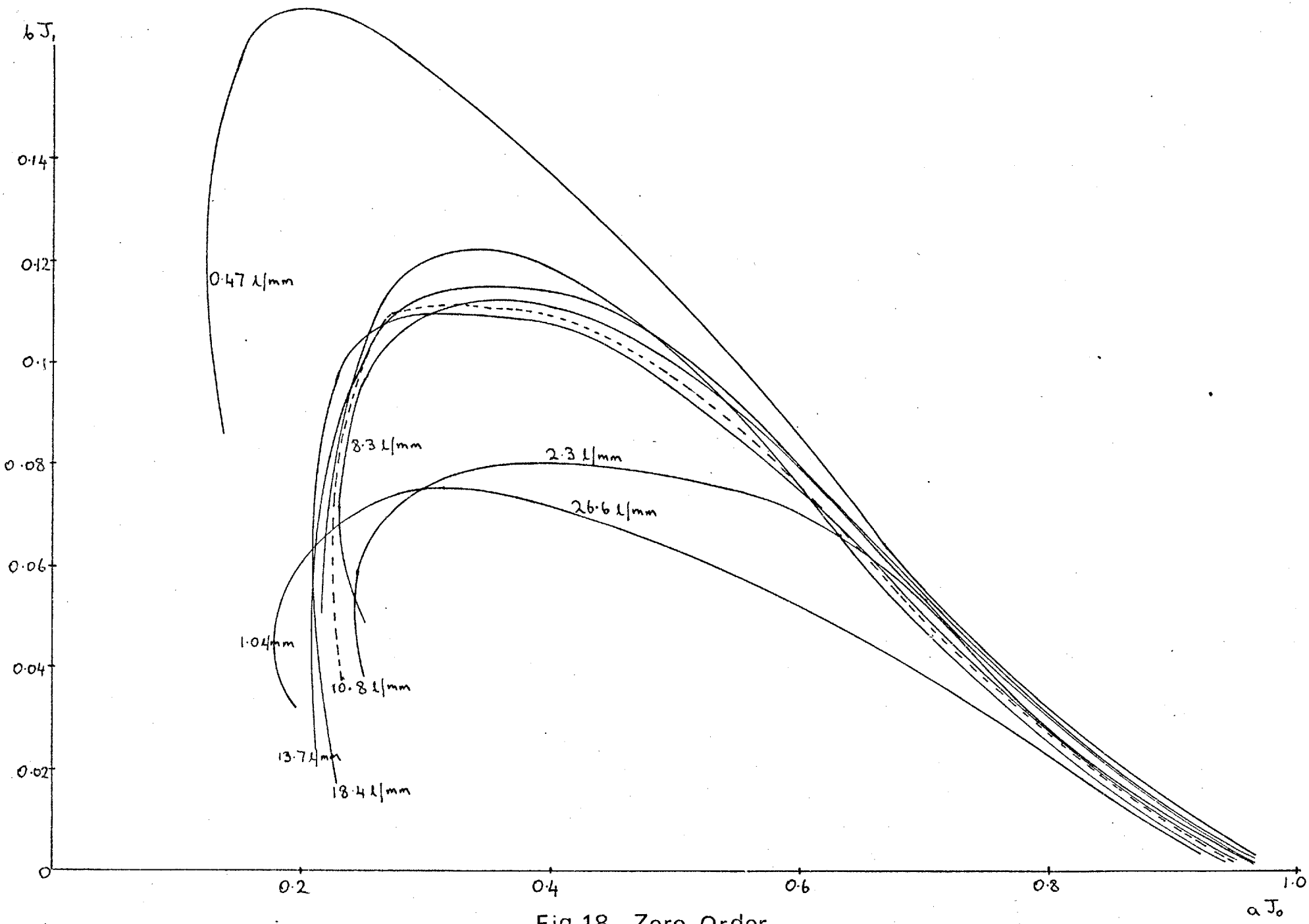


Fig 18 Zero Order

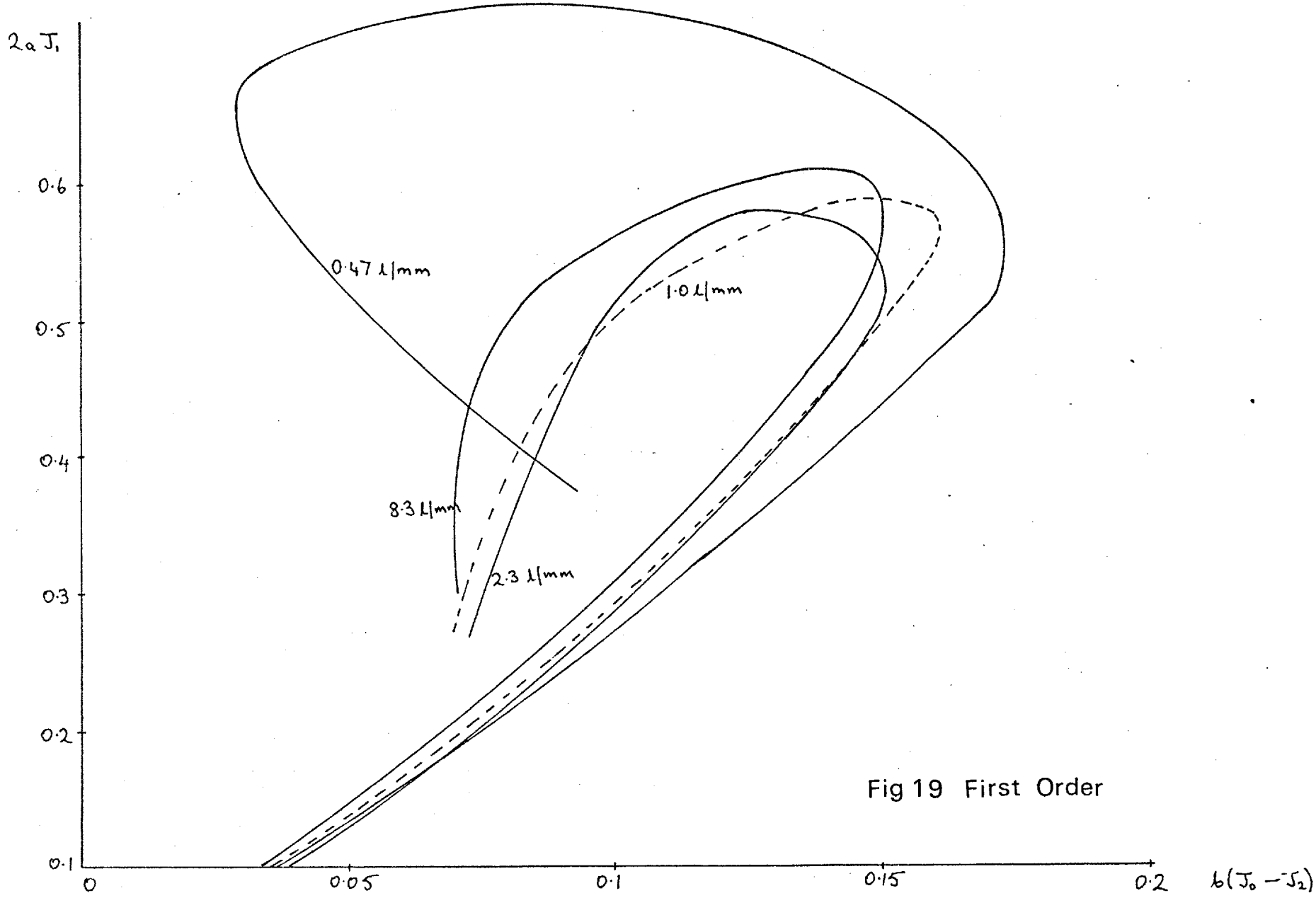


Fig 19 First Order

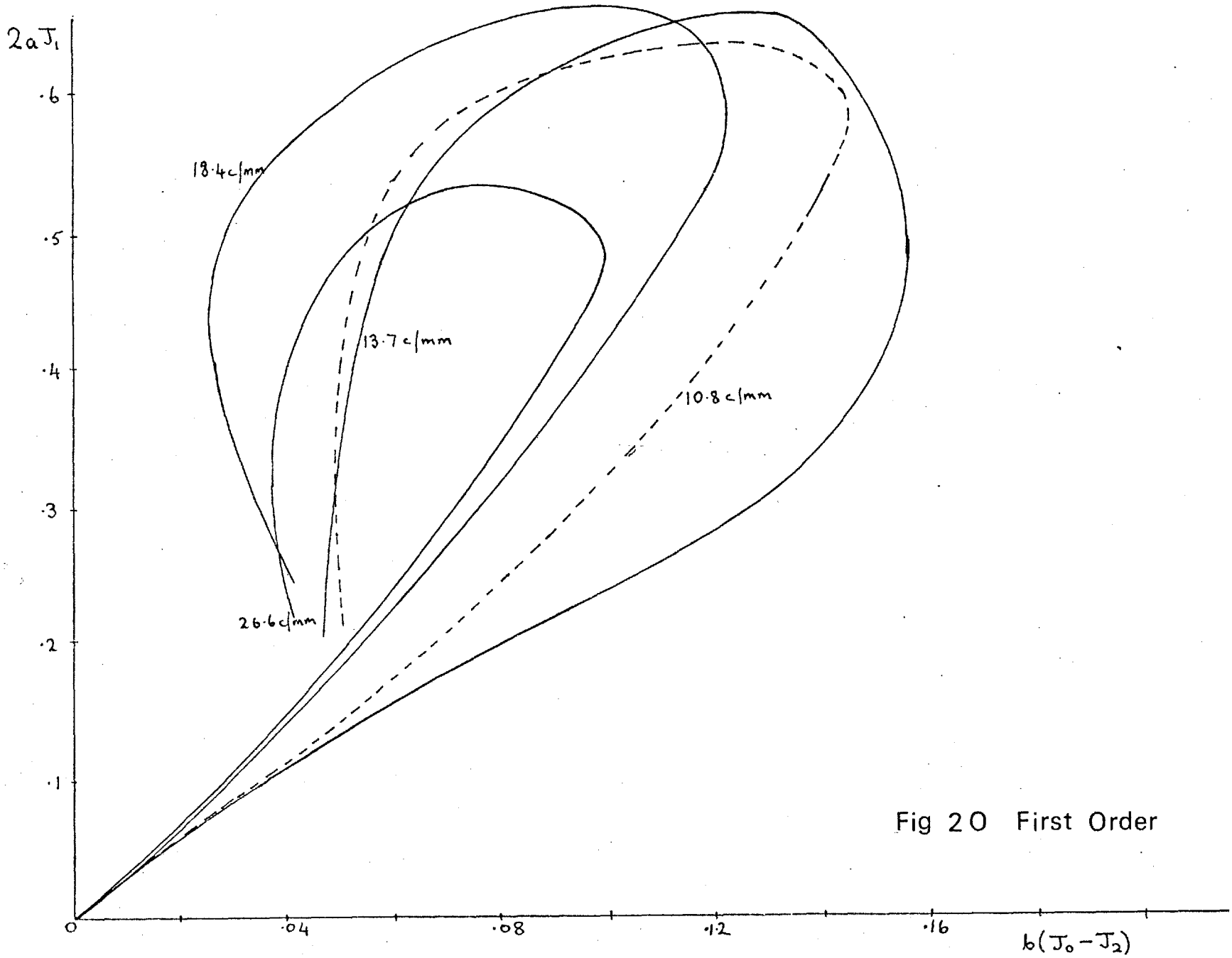


Fig 20 First Order

but simply rotates the curves about the origin of the complex plane. The value of z in the region where the curves turn over is of the order of one. The value of b, the amplitude of the sine wave, was kept constant as far as possible. The value of a, the bias amplitude transmittance, was varied from 0.9 down to about 0.3. At the 0.9 value of a the value of b was somewhat lower than for the other values. Also due to the limitations of the emulsion the value of b decreased slightly with the higher spatial frequencies.

Fig. 18 shows the variation of the zero order spectrum. From these curves it is seen that the amplitude of the zero order decreases in the same fashion for the different frequencies; further, with the exception of the curves at 0.47c/mm and 26.6c/mm all the other curves follow virtually the same values. The value of the zero order decreases fairly linearly from one (clear emulsion) and then turns over and begins to go in the shape of the spiral as shown in chapter II.7. The value at which the curves turn over occurs at an amplitude transmittance of about 0.6 at which point it has an amplitude value of 0.3 giving an intensity value of 0.09.

Of more interest is the amplitude of the first order spectra, since this would be used to identify any frequency structure. The curves for all the different frequencies examined, (Figs. 19, 20) are of the same nature. As the bias amplitude decreases so the amount of light in the zero order spectrum increases. The maximum amplitude for the first order spectrum occurs at a bias amplitude transmittance value of about 0.7. The amplitude vector then turns over and decreases in a similar manner. The bias

amplitude transmittance for maximum vector amplitude coincides with the centre of the linear portion of the T_a - E curve, (Fig. 7) and this is the expected result. As regards increasing the phase retardation constant it is seen that it can enhance the first order amplitude. The exception to this appears to be for the two extreme frequencies examined. At the low frequency, 0.47c/mm, the value of the first order spectra is rather high, whereas that of the high frequency 26.6c/mm is rather low. The low frequency discrepancy is not very significant. Its probable cause is in measuring the phase retardation constant. We have measured the retardation in transmission through the emulsion. This means that the phase variations of the base are superimposed on relief retardations. At high frequencies the relief phase retardations changes are much more rapid than those of the base and hence we have no interference from the base. At a frequency of 0.47c/mm we are approaching the rate of the phase fluctuations of the base, and hence it is possible that we have added on some of the base changes to the relief phase retardations. This would give a higher value for the relief phase retardation constant which would account for the large amplitude of the first order diffracted spectra. In any case, it is not very realistic to talk about the amplitude of the first order spectra of such low frequencies because they occur so close the centre of the zero order. This makes it extremely difficult, if not impossible, to isolate them, added to the fact that they would be enhanced by high low angle scatter.

The low value at the higher frequency can be explained by considering the performance of the emulsion at this frequency.

Although we have been able to record the correct bias amplitude the modulation amplitude has decreased and as seen from Fig. 7, Ch. II, this results in a decreased first order amplitude.

From Figs. 19 and 20 we see that the maximum total amplitude of the light diffracted into both the first orders is on average about 0.6. This gives an amplitude to each order of 0.3 and an intensity of 0.09. This of course means a diffraction efficiency of 9%. Hence we have again demonstrated that the relief phase variation can enhance the diffraction efficiency. The value of the amplitude of the sinusoidal modulations for Figs. 19 and 20 was around 0.18; this is a little below the optimum value of 0.23 obtained from Fig. 7, Ch. II and hence theoretically, we can expect to improve a little on this efficiency.

So far we have looked at the effect of changing the amplitude of the modulations and the bias amplitude transmittance on the first order spectra. We have one more variable, the phase retardation constant, which we must take into consideration. If we could keep both a and b constant and increase the phase modulation we would then have the complete solution. Fig. 21 shows how the first order amplitude vector changes with increased phase modulation. The curve has been plotted with $b/a = 0.2/0.7$ and k/λ increasing from zero to four. From this curve it is seen that the maximum diffraction efficiency would occur when k has a value of λ. At this point we have the maximum diffraction efficiency that we can expect using this type of grating. The amplitude of one of the spectra is 0.4 giving a maximum diffraction efficiency of 16%.

We have seen that the phase retardation produced by density

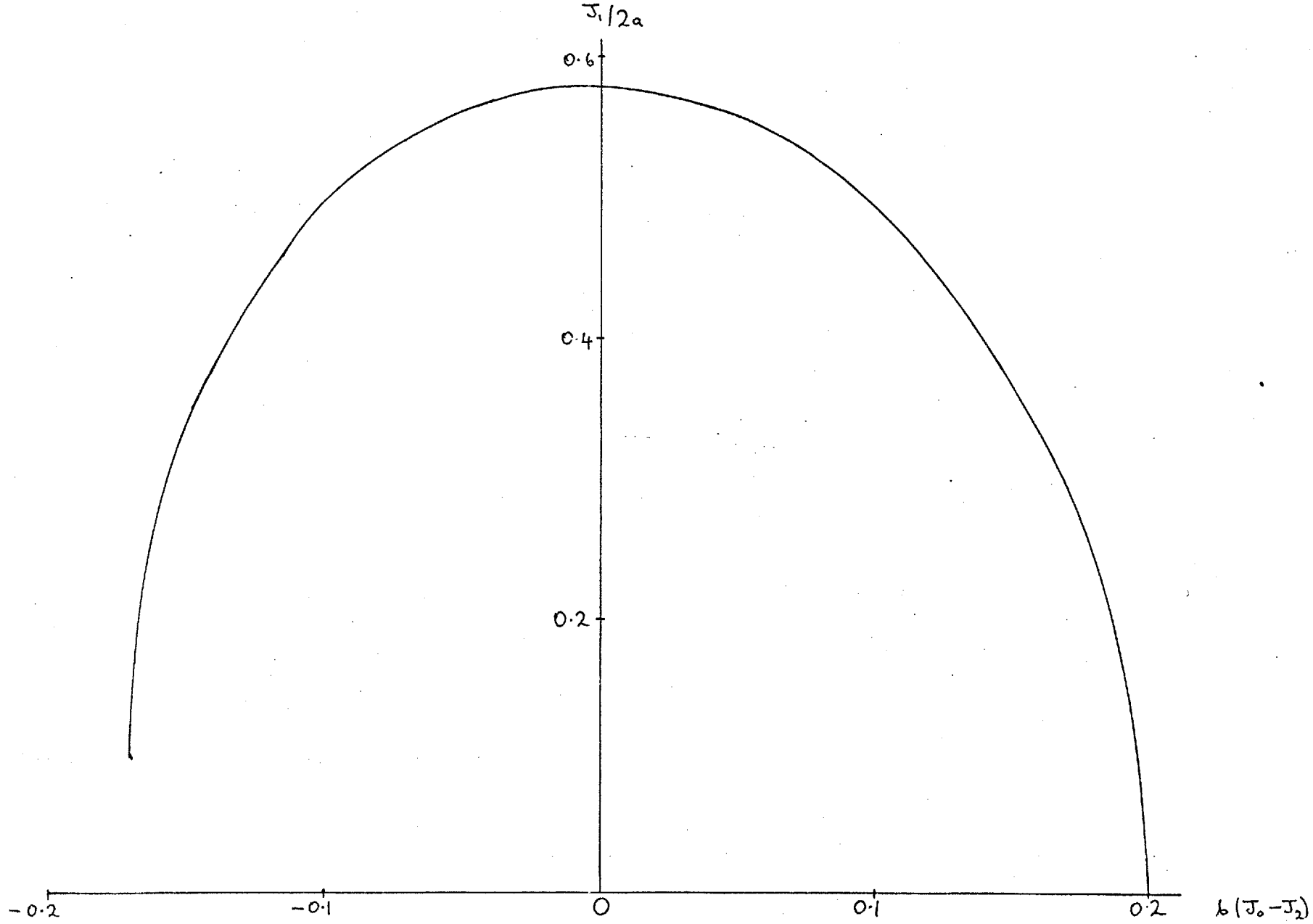


Fig 21 First Order

variations can enhance the first order spectra and thus the diffraction efficiency. For this particular emulsion the optimum conditions occur when b = 0.23, a = 0.7 and k = λ ; this would give us a diffraction efficiency of 16%. In practice however, this can never be expected, mainly because we are not in a position to be able to control the three variables to a sufficient accuracy. Furthermore it is not expected that a value of k = λ can be achieved with this particular emulsion and phentrace developer. This would limit the maximum diffraction efficiency to about 14%. For this type of grating therefore, we can possibly expect efficiencies in excess of the 6.25% predicted for pure amplitude gratings.

V.5. Diffraction Efficiency

In section II.5 we have defined diffraction efficiency as the ratio of the diffracted flux in one of the first order spectra to the flux through a clear portion of the emulsion. In the previous section we have discussed how the amplitude of the diffracted spectra varies for the different recording signals. In the present section we will show that the values of the amplitude of modulation and the bias amplitude transmittance giving optimum diffraction efficiency are in agreement with those expected. The absolute values of the diffraction efficiency are however lower than the maximum predicted by the theory. This is to be expected since we know that the photographic emulsion is not a perfect recording medium.

Since we are limited to the developers we are able to use, it is not possible to vary the phase retardation constant at a given spatial frequency. We can of course change the constant by changing

the spatial frequency, but since it is expected that the emulsion performance changes with frequency, it would not be sensible to use such a method. We are therefore limited to varying the amplitude modulation and bias amplitude transmittance and to observe their effects on the diffraction efficiency. From Fig. 7, Ch. II it is see that the effect of changing the amplitude at a given bias transmittance (0.5) is least at around the optimum value of 0.23. As we increase the modulation from 0.2 through 0.23 up to 0.27 the diffraction efficiency goes from 7.9% up to 8.4% then back to 7.9%, which is a total change of only about 6% of its value. Since the experimental accuracy of measuring the diffraction efficiency is of this order we cannot sensibly investigate the modulation effect around its optimum value. The value of the bias amplitude transmittance was varied from 0.8 to 0.6 at a spatial frequency of 7.5c/mm with the amplitude modulation remaining sensibly constant. Fig. 22 shows the change of diffraction efficiency within this range for both theoretical and experimentally observed values. From this curve it is seen that the observed values are about a factor of three down on the theoretical values in the bias amplitude transmittance range 0.8 to 0.7. As the bias is lowered to 0.6 the observed values drop far more sharply than predicted, which is of course expected. When one considers the linear portion of the Ta - E curve it is seen that it extends from 0.9 to just about 0.5. The modulation of the recorded signal has an amplitude around 0.23, which means that once we lower the bias transmittance below 0.7 we are getting into a very non-linear region. This has the effect of introducing harmonics which give rise to higher order spectra with

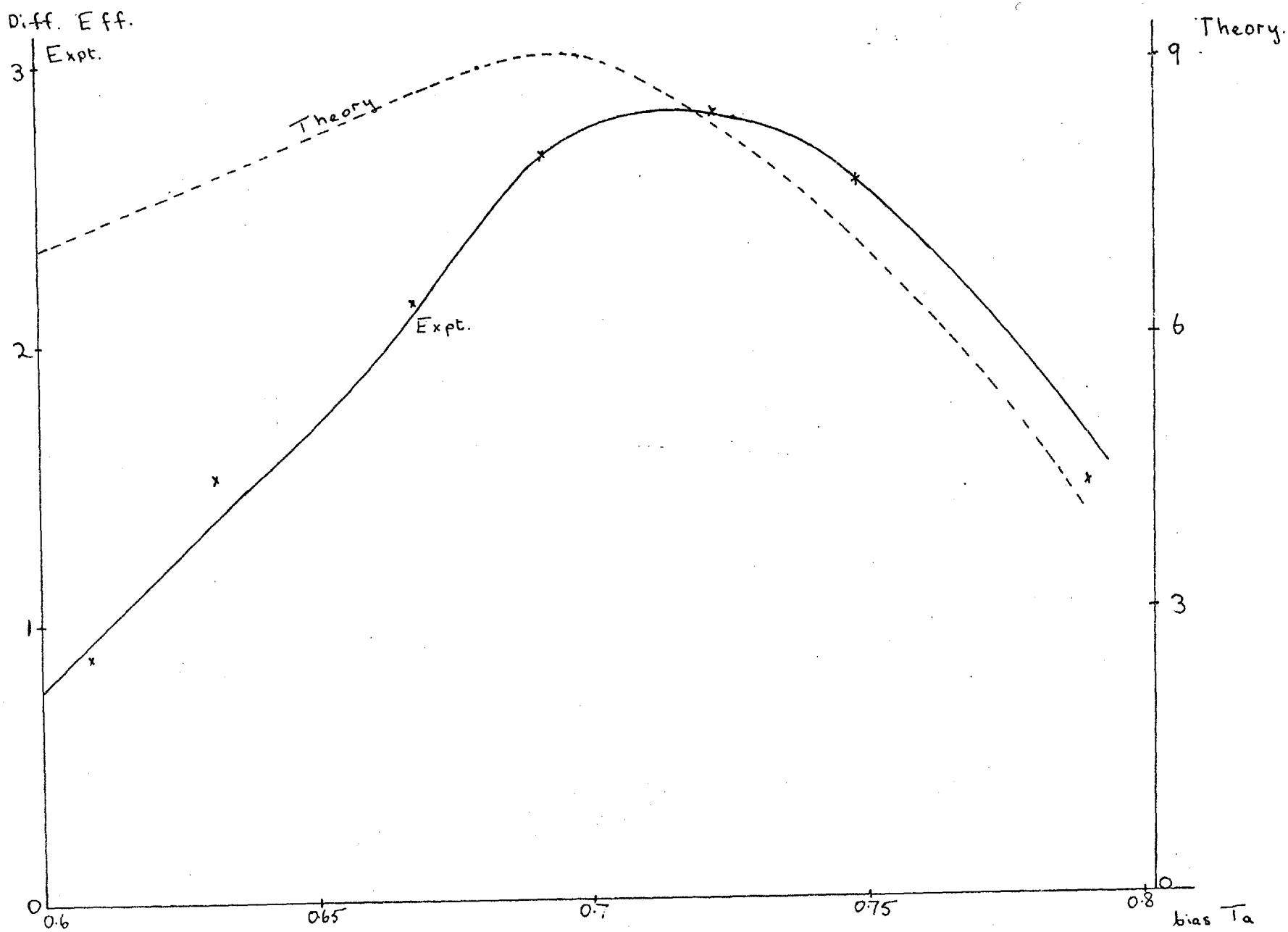


Fig 22 Diffraction Efficiency

an overall loss of diffraction efficiency. The zero order was also measured and found to agree reasonably well with theory; its value lay between 20 and 30%.

The diffraction efficiency was also checked at two other spatial frequencies, 12.5c/mm and 24c/mm. Theoretically the efficiency should now increase to around 13% and 14% respectively for optimum recording. The maximum diffraction efficiency observed at 12.5c/mm did increase 3.3% but at 24c/mm it had dropped to only 1.7%. These results indicate that at 12.5c/mm the emulsion performance is still good and the phase retardation tends to increase the diffraction efficiency. At 24c/mm the emulsion performance has dropped and it is not possible to record near the optimum conditions, hence we have a reduced efficiency.

So far we have been considering the diffraction efficiency of the emulsion when it is being used as an amplitude grating with phase effects superimposed. We also know that we can expect higher efficiences from a pure phase grating than an amplitude grating. In order to obtain a phase grating we need to bleach an amplitude grating. To use the present emulsion would however be defeating its purpose since the relatively long time of the bleaching process would mean that we no longer had a rapid access emulsion. We also suspect, V.4, that this emulsion is not particularly suitable for bleaching. It was nevertheless decided to obtain an idea of the diffraction efficiency of this bleached emulsion.

The bleached image with known phase retardation was obtained as described in V.4. The magnitude of the zero and first order spectra were then measured for various phase retardations and various

original amplitude transmittance values. The value of the zero order remained constant to within 6% of its value for one particular frequency and within 12% for the whole range of frequencies. The value of the diffraction efficiency was found to be dependent on phase retardation and also on the initial recording conditions. This result indicates that non-linear amplitude recording produces a non-linear phase relationship, as is reasonable to expect. As a comparison, the diffraction efficiency for the same value of phase retardation at the same frequency dropped from 4.2% when initially recorded on the linear region to 1.3% when recorded in the non-linear region. The optimum value of the diffraction efficiency is obtained by optimising on the initial recording conditions and phase retardations at the same time. As a further example it was seen that at one frequency when the phase retardation increased from 0.23 λ to 0.31 λ and the bias amplitude transmittance dropped from the linear region of 0.68 to the non-linear region of 0.5 the diffraction efficiency stayed constant within the experimental accuracy. The maximum diffraction efficiency observed with this emulsion was of the order of 5.5%. Hence it is seen that even though we do not have a particularly suitable bleaching process, nevertheless we can still obtain a higher diffraction efficiency from the phase grating than the corresponding amplitude grating.

V.6. Effect of Base Phase Variations on Diffraction Plane

In V.3 we investigated the phase properties of the base and the emulsion. Our first conclusion was that most of the phase variations present were due to the base itself. We then imposed

a tolerance criterion on the flatness of the film and found a relationship between the percentage of areas within the tolerance as a function of area. In this section we will investigate the effect of the phase variations present in the base on the distribution of light intensity in the diffraction plane. The main reasons for not using the emulsion in these investigations were firstly, the emulsion has only a little effect on the phase variations, and secondly, the effect of scatter produced by the emulsion would interfere with the results.

To perform the experiments the coherent processor was set up in conjunction with the Mach-Zehnder interferometer as described in IV.2. In this configuration we are able to measure both the distribution of light intensity in the diffraction plane and also the variance of the wavefront shape. Three different size apertures were used, these being 1.5cm, 1.0cm and 0.7cm radius circular apertures. Thus we were examining the effect of the phase variations due to 175sq.mm, 78sq.mm and 38sq.mm of the base. From the previous experiments of V.3 we know that the percentage of areas within the tolerance limit for these area sizes are 6%, 15%, and 26%. For these experiments we examined eight 175sq.mm areas, six 78sq.mm areas and six 38sq.mm areas. If our probabilities are correct we would expect a total of about two or three areas that would satisfy our tolerance criterion. Besides the phase distribution of the base, there are two other factors which must be taken into consideration when examining the light distribution. The first of these is the contribution from the rest of the optical system. To allow for this we first of all scanned the diffraction plane,

with no emulsion present, for the three apertures. These results can be seen in Figs. 23, 32 and 39. From Fig. 39 we can see that for the 0.7cm aperture we are effectively diffraction limited. Fig. 32 shows the light intensity distribution for the 1.0cm aperture, here again we are near the diffraction limit. For the 1.5cm aperture, (Fig. 23), we can see that the intensity of the first maximum is in excess of 2% of the central peak, which is slightly over the theoretical (1.7%). Hence we only need to take into consideration the rest of the optical system when we are using the 1.5cm aperture. It is more difficult to make an allowance for the second factor which should be taken into consideration. This is the scattering from scratches and possible impurities in the base, the effect of the scratches being the more prominent. The scratches occur when the film is being pulled across a camera or through a holder. The direction of the scratches is therefore along the length of the film and thus the scattering would be shown up in the scans across the emulsion. The magnitude of their effect could not easily be determined since it would be dependent on several factors such as the number and depth of scratches. Their presence could be used to account for an unusually high light intensity outside the central maximum in one direction when this is not indicated by the variance of the wavefront. Such a diffraction image is typical of astigmatism when viewed in one of the focal planes. It is however unlikely to be the cause of our image since its presence is not indicated by the variance and in any case such an astigmatic image cannot be expected for rays passing through such a thin piece of material.

In eqn. V.1 we gave a relationship between the intensity at the centre of the diffraction pattern and the variance of the wavefront. In deriving this equation[8] it was assumed that the aberrations were small so that third and higher powers were neglected. The approximation is good for intensities within the Strehl limit, but becomes poor when the Strehl ratio drops below 0.7. It is therefore not meaningful to obtain the Strehl ratio from the variance of the wavefront for most of our cases, although it is expected that the Strehl ratio would decrease with increasing variance. Furthermore due to the random nature of the wavefront deformation it is not possible to correlate the irregularities introduced with the Seidel aberrations.

In principle the spread function could be calculated in the usual way as a Fourier transform; this however would not yield any useful correlations since the processor already gives the transform and the object of using the Strehl ratio is to get a single number for performance. Shannon[70] shows the effect of a small random wavefront error on the spread function and also the transfer function for the same errors. He states that visually the amount of random wavefront aberration can be seen to have a consistent effect upon the transfer function, but he does not derive any direct mathematical correlation. In general, all we can do is to compute the modulation transfer function from the wavefront shape. Visual inspection of the spread function gives a good indication of the modulation transfer function, all we will therefore do is to simply describe some of the effects produced by the base deformations.

In the following figures scans along the emulsion are given by the broken line and those across by the continuous line. The two curves for each direction represent the scan on either side of the central maximum.

(a) 1.5cm diameter aperture, Figs. 23 - 31

Of the eight portions of base examined at this aperture, none were found to satisfy the Strehl criterion. The lowest variance for this series was $0.02\lambda^2$ which corresponds to Fig. 27; the highest variance was $0.12\lambda^2$ which gave rise to Figs. 24 and 29.

(i) Fig. 24. As regards phase deformation, this portion of base was the worse of the whole of the present series. Scans across the emulsion produced a first peak intensity of nearly 60% of the central maximum and also a secondary maximum occurs at nearly 9% intensity. The scan along the emulsion shows a first maximum of over 4% on one side and no well defined maxima at all on the other side. It is possible that scratches are responsible for some of the light found in the secondary maxima in the scan across the emulsion. The positions of the first minima occur too close to the central peak.

(ii) Fig. 25. Positions of the maxima and minima agree well with the normalised values for scans along the emulsion. The values of the light intensity appear to be a little lower than expected. The scan across the emulsion shows one peak at 30% but no well defined peak on the other side. Overall it is seen that most of the energy is contained in the central maximum; the value of the variance is $0.03\lambda^2$.

(iii)Fig. 26. Although most of the light is within the central

maximum, its area is much larger than expected. The highest value of the first maximum is 12%, the variance for this base being $0.04\lambda^2$.

(iv) Fig. 27. This turns out to be the best of this size base as regards both variance and intensity distribution. The appearance of the spread function gives a slight indication of coma.

(v) Fig. 28. As regards variance this base has approximately the same value as that of Fig. 26. The shape of the spread function is however completely different. The intensities of the first maxima are much greater than those of Fig. 26, but they occur much closer to the central maximum. The resultant effect is that the amount of energy outside the theoretical central disc is of the same order of magnitude.

(vi) Fig. 29. Has the same variance as Fig. 24 but much less light outside the central maximum. As previously explained this could be accounted for by the existence of scratches.

(vii) Fig. 30. In a way, similar to Fig. 25 but more spread out, resulting in a higher wavefront variance of $0.08\ \lambda^2$.

(viii) Fig. 31. Shows reasonable high secondary maxima for scans in both directions. The positions of the maxima occur a little before the theoretical values. The variance for this base was found to be $0.05\ \lambda^2$.

(b) 1.0 cm Diameter Aperture

The variance of the wavefront for this area base was a little lower than the larger size. One portion of the base was found to be within the Strehl limit, the others had a variance between $0.02\lambda^2$ and $0.03\lambda^2$.

(i) Fig. 33. The scan along the length of the emulsion seems to

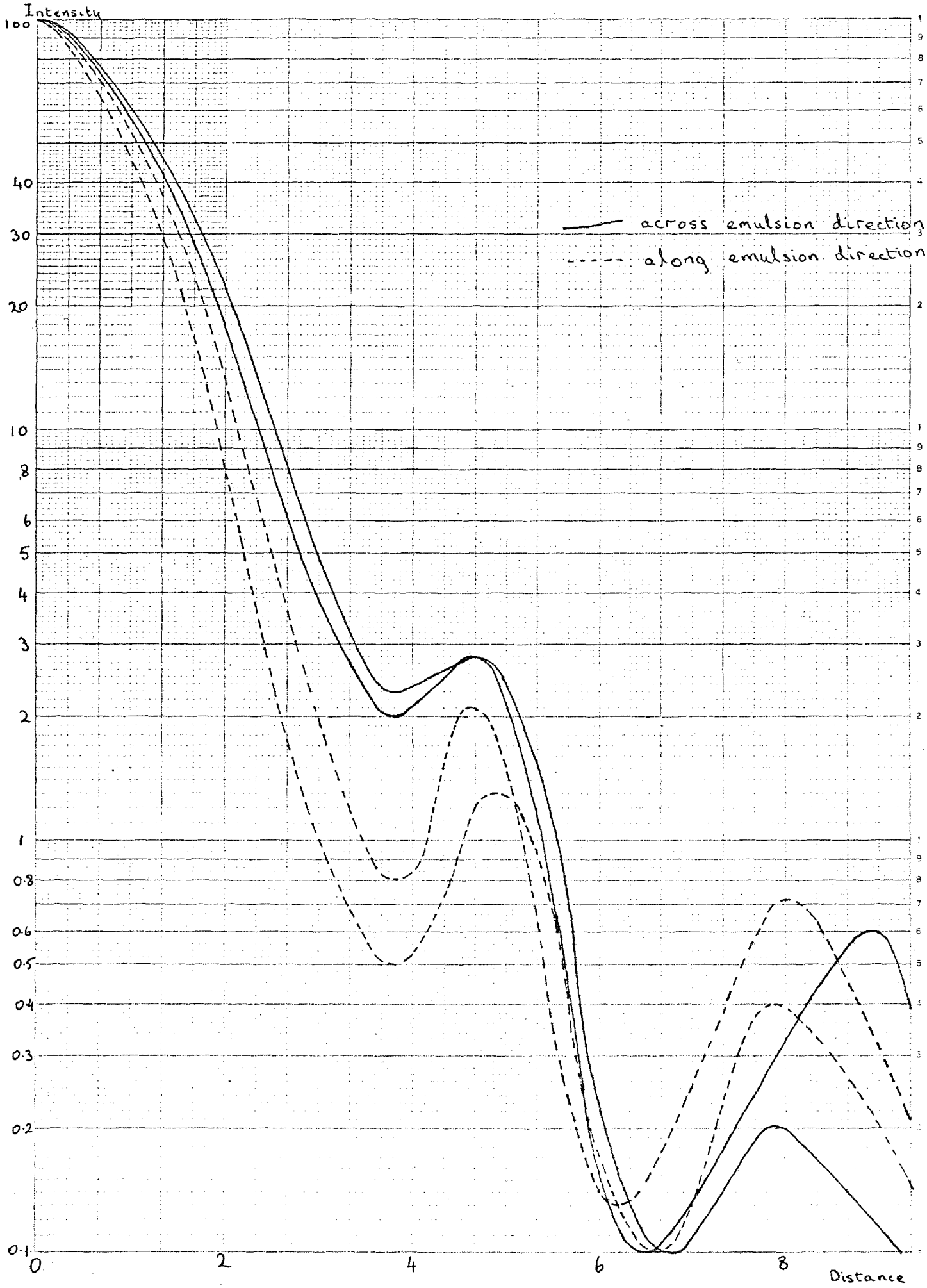


Fig 23. No Emulsion 175 sq. mm. apt.

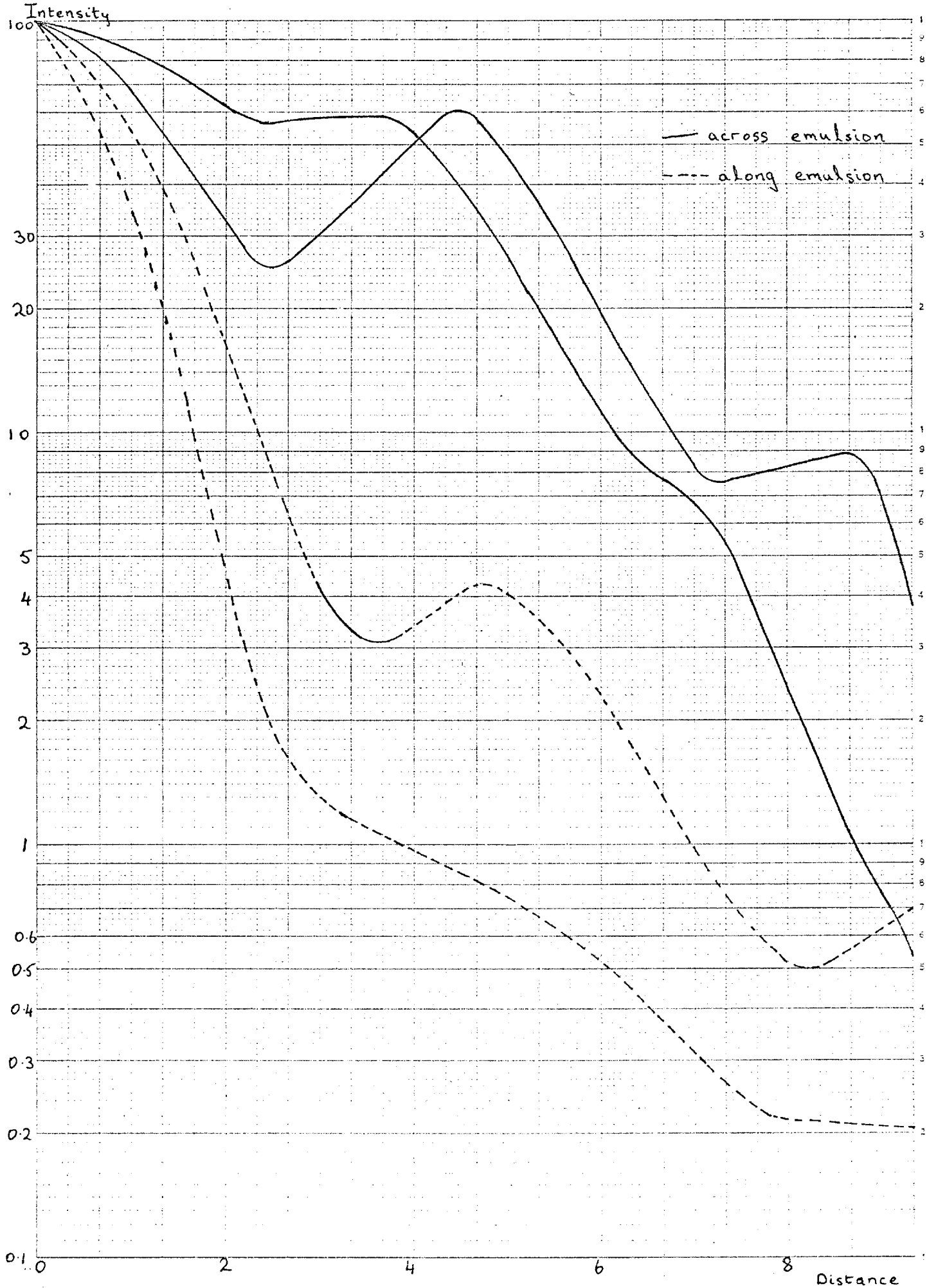


Fig 24 175 sq. mm apt.

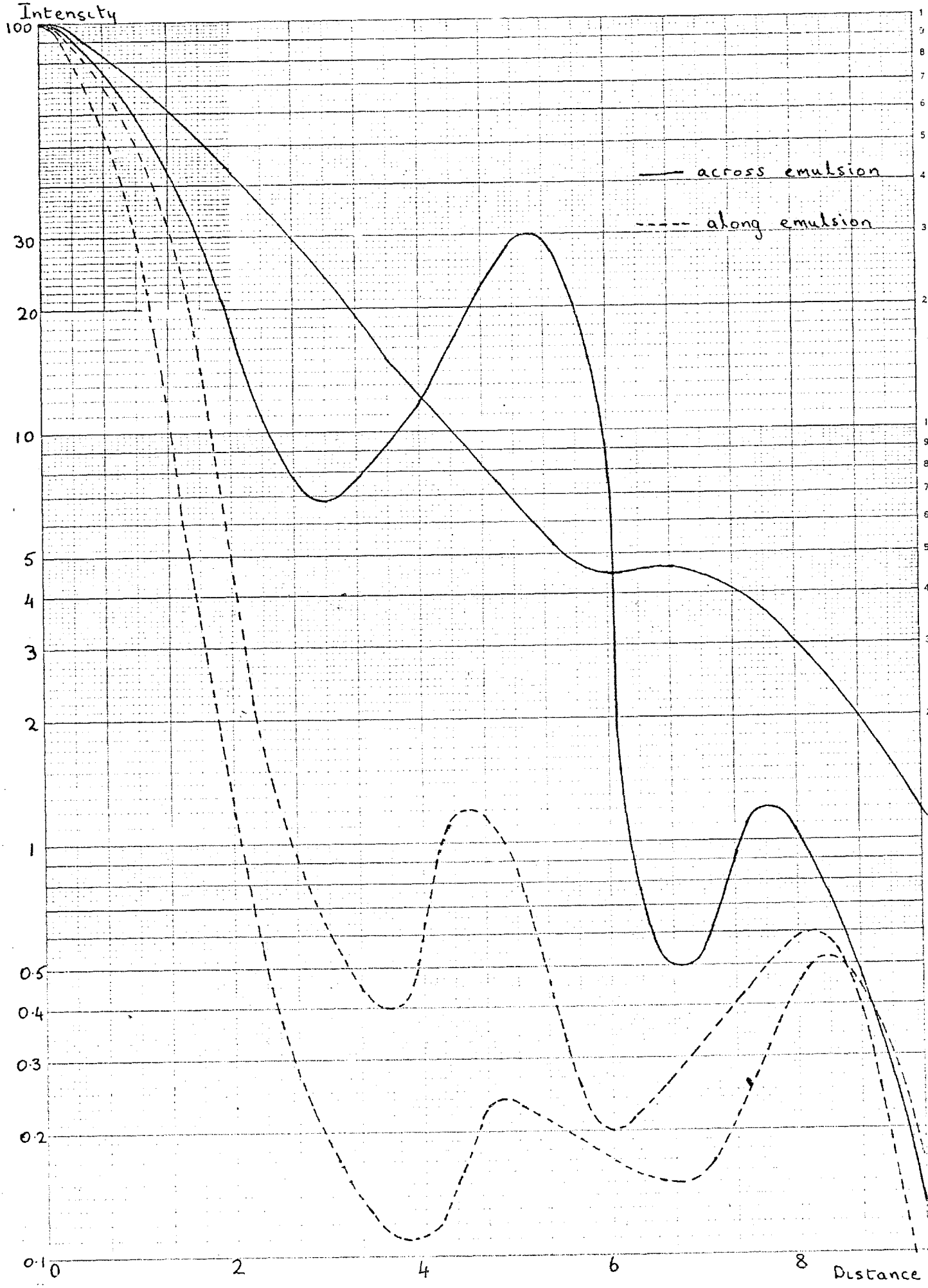


Fig 25 175 sq. mm. apt.

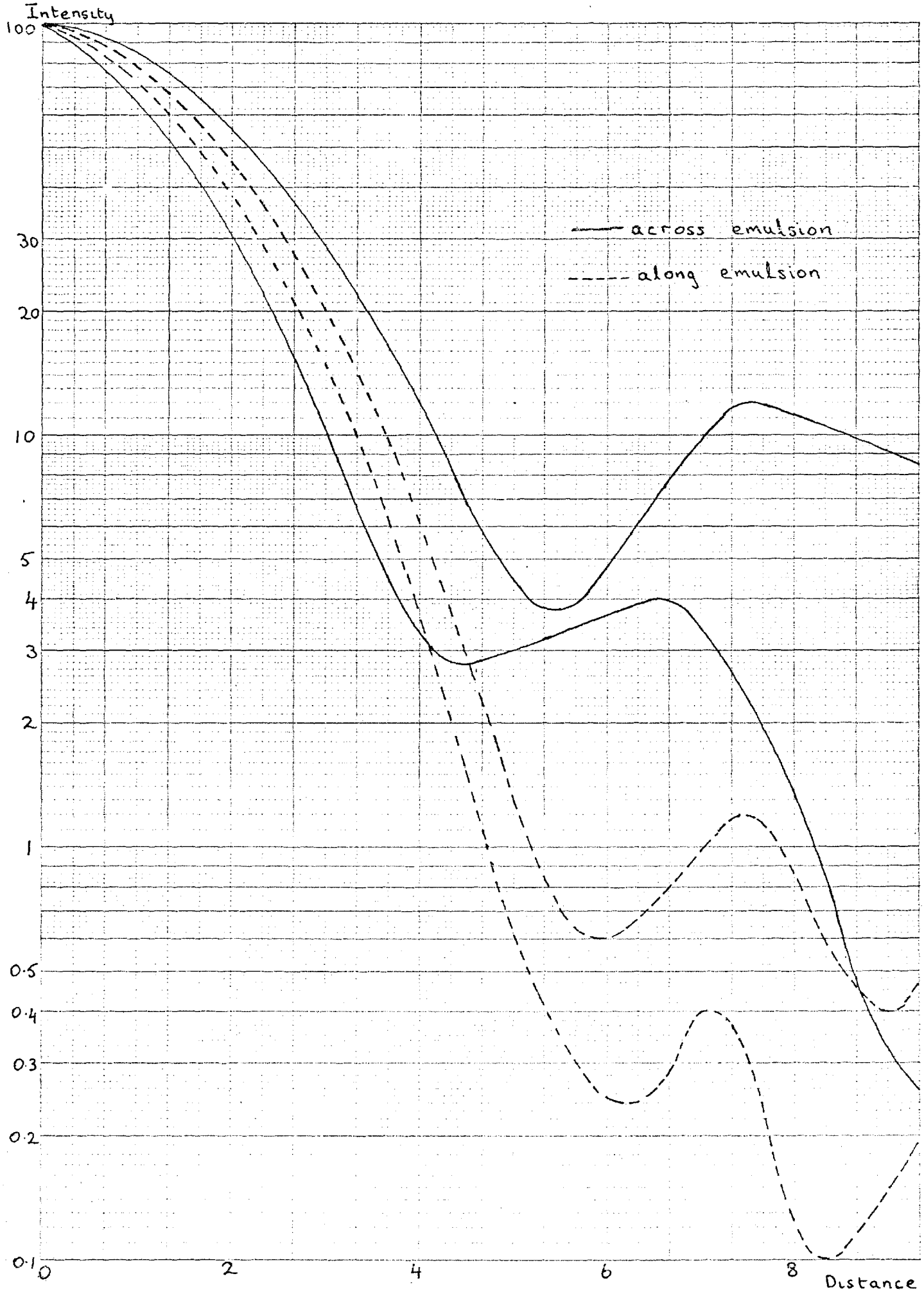


Fig 26 175 sq. mm. apt.

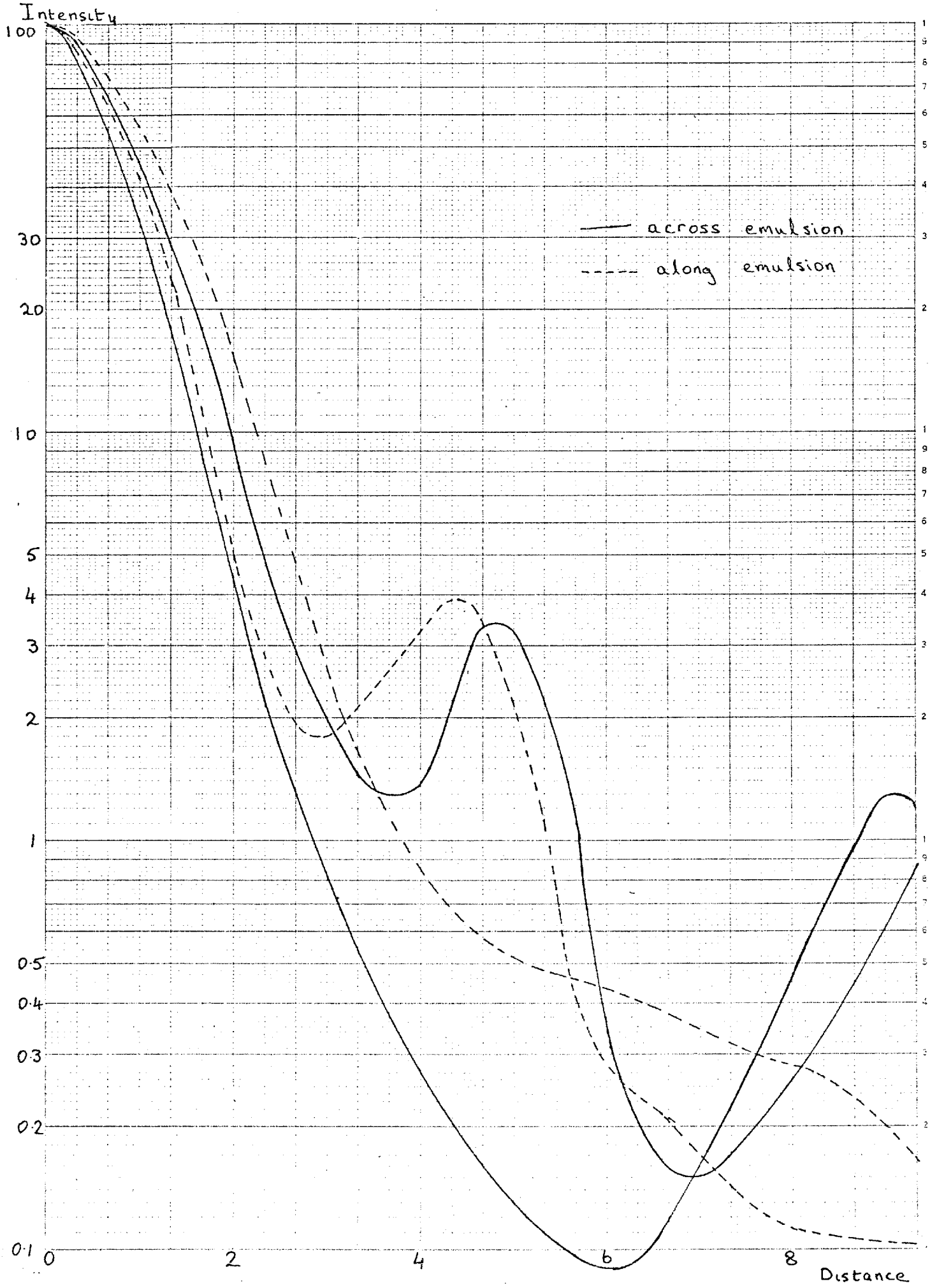


Fig 27 175 sq. mm apt.

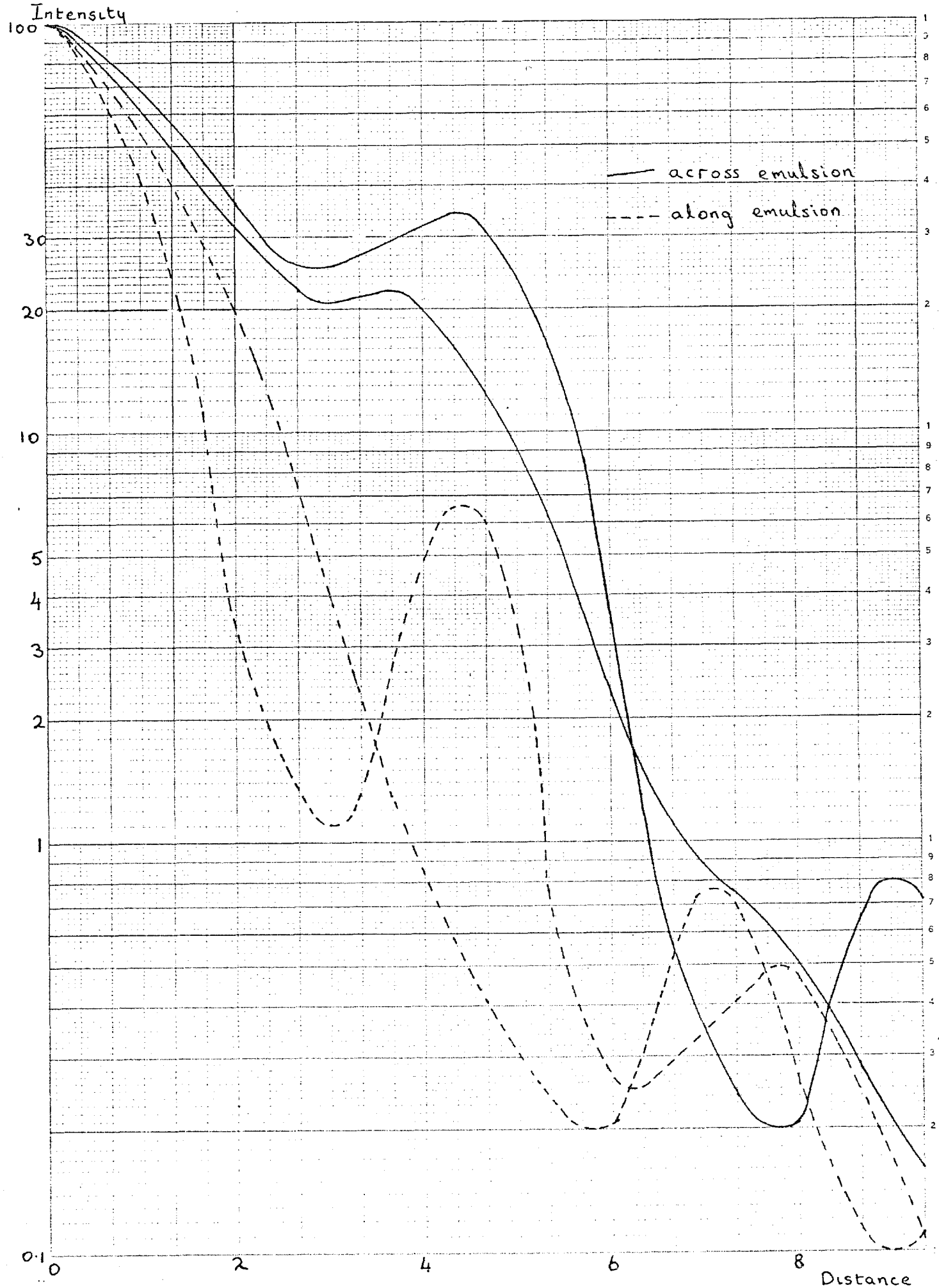


Fig 28 175 sq. mm apt.

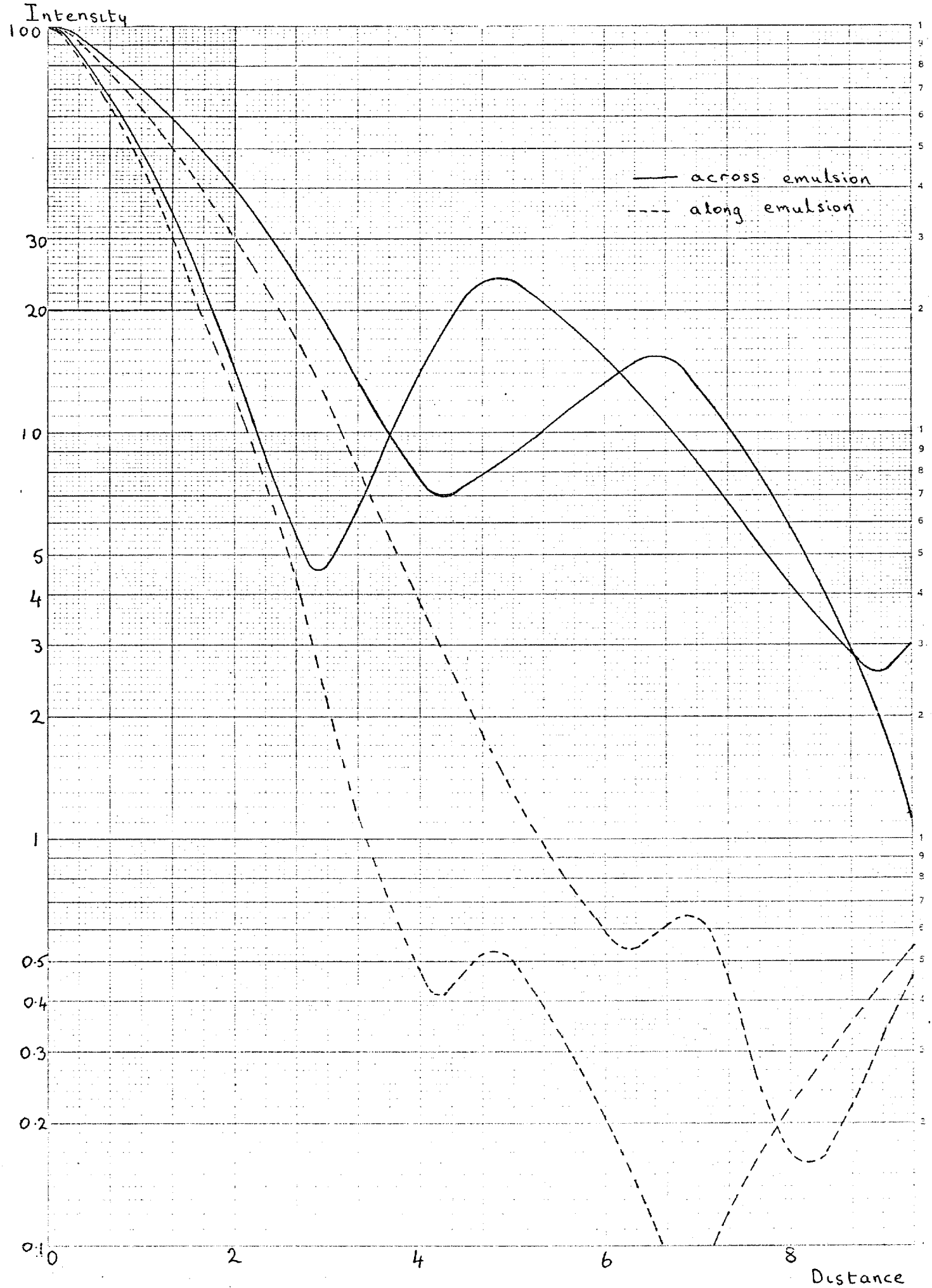


Fig 29 175 sq. mm. apt.

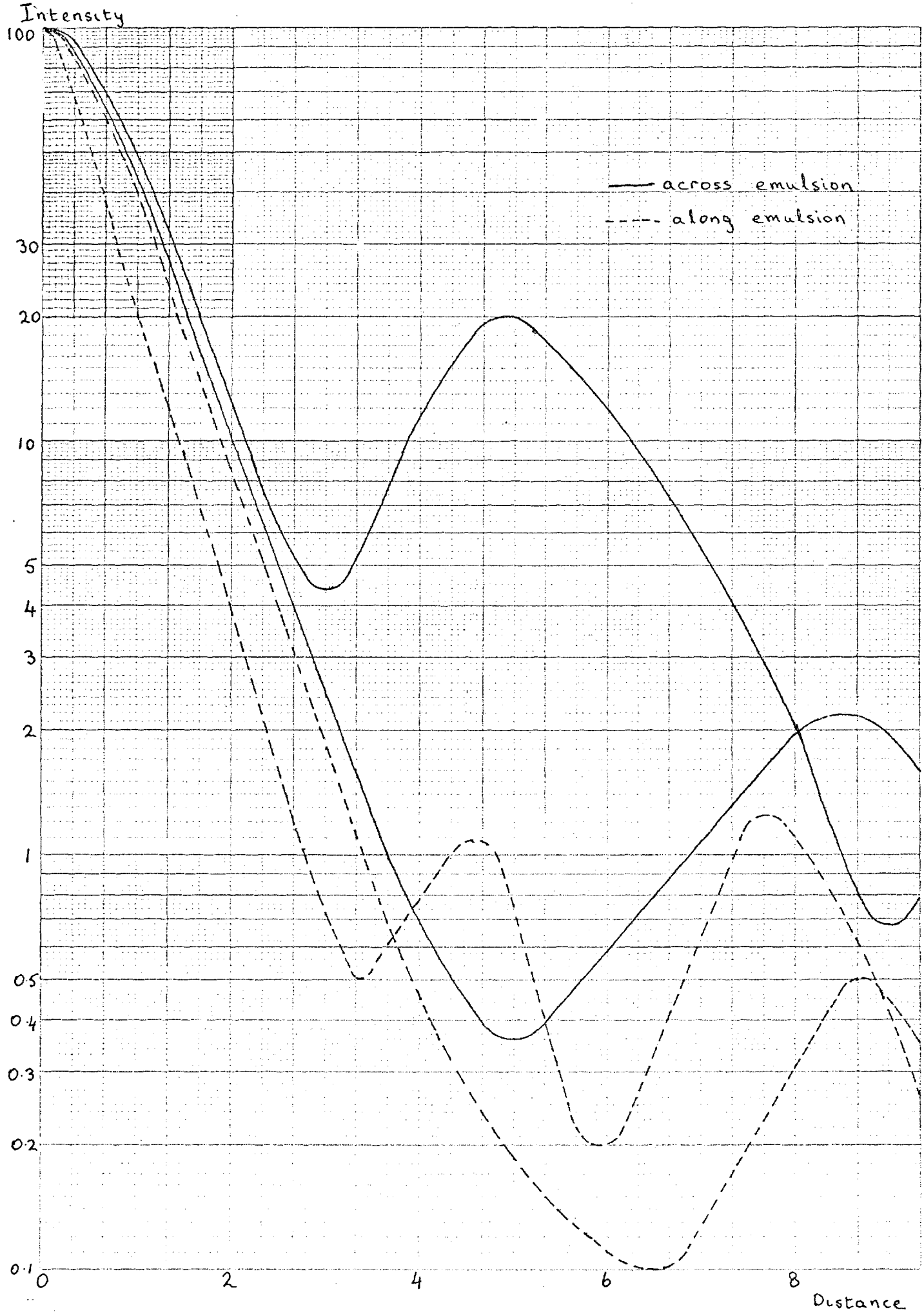


Fig 30 175 sq. mm apt.

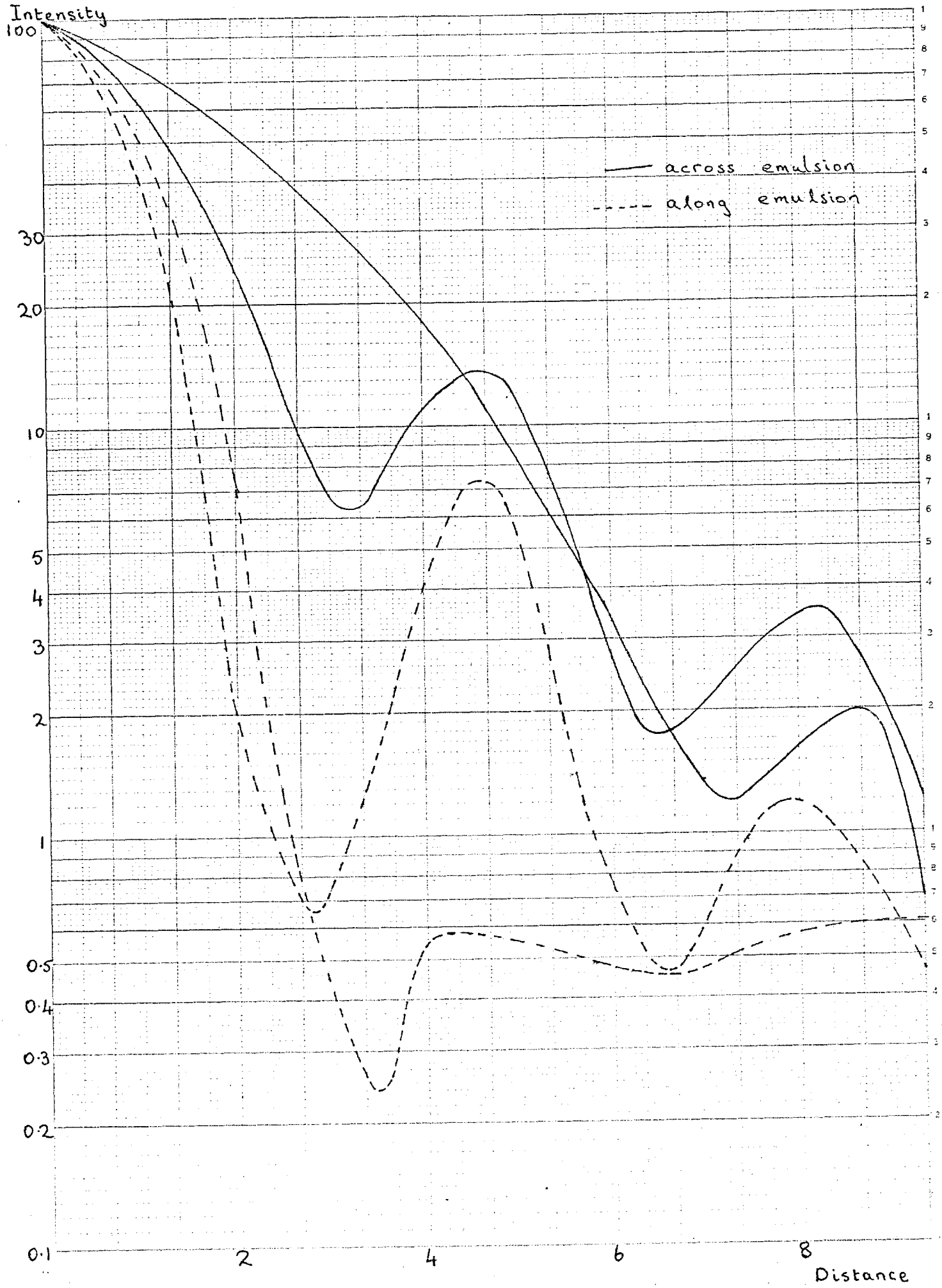


Fig 31. 175 sq mm apt.

indicate a reasonably good portion of base. The scan across the emulsion however showed no definite first maxima or minima. The reason for this could again be due to scratches, since the value of the variance, $0.03\lambda^2$ did not indicate this.

(ii) Fig. 34. Here again the scan along the emulsion agreed well with that obtained without the film. The scan across the base gave one relatively high peak at 16% as was expected from the value of the variance, this being $0.02\lambda^2$.

(iii) Fig. 35. The value of the variance was of the same order of magnitude as the previous figure. The diffraction pattern obtained however, showed reasonably high peaks for both the scans with no clearly defined first minima. It is expected therefore that the majority of the contribution in the high orders is due to scattering from scratches.

(iv) Fig. 36. For this portion of base, the variance of the wavefront indicated an acceptable area ($0.005\lambda^2$). This was confirmed in the scans in both directions which agreed well with the expected value.

(v) Fig. 37. The presence of a high peak, over 30%, for the scan across the base, indicated again the influence of scratches. The scan along the emulsion was more consistent with the variance value of $0.02\lambda^2$.

(vi) Fig. 38. The variance for this portion was of the same order of magnitude as for most of the other figures. The spread function however, showed no clearly defined first maxima or minima for either of the scans, just a shoulder occuring in approximately the correct position. The value of this shoulder was rather high and it is thought that some form of scattering effect other than scratches was present.

(c) 0.7 cm Diameter Aperture

For these areas one portion of base was found to have a variance

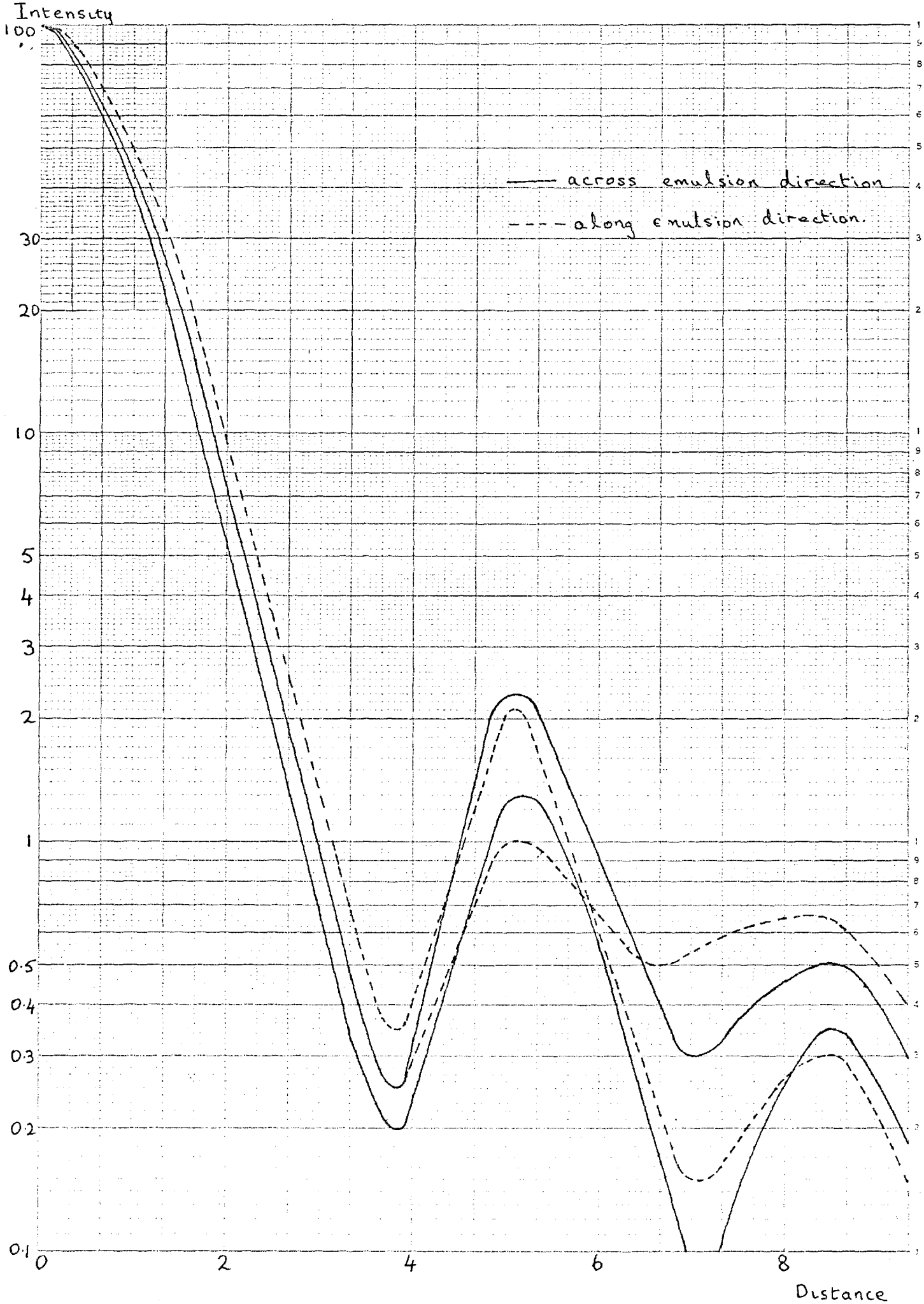


Fig 32 No Emulsion 78 sq. mm. apt.

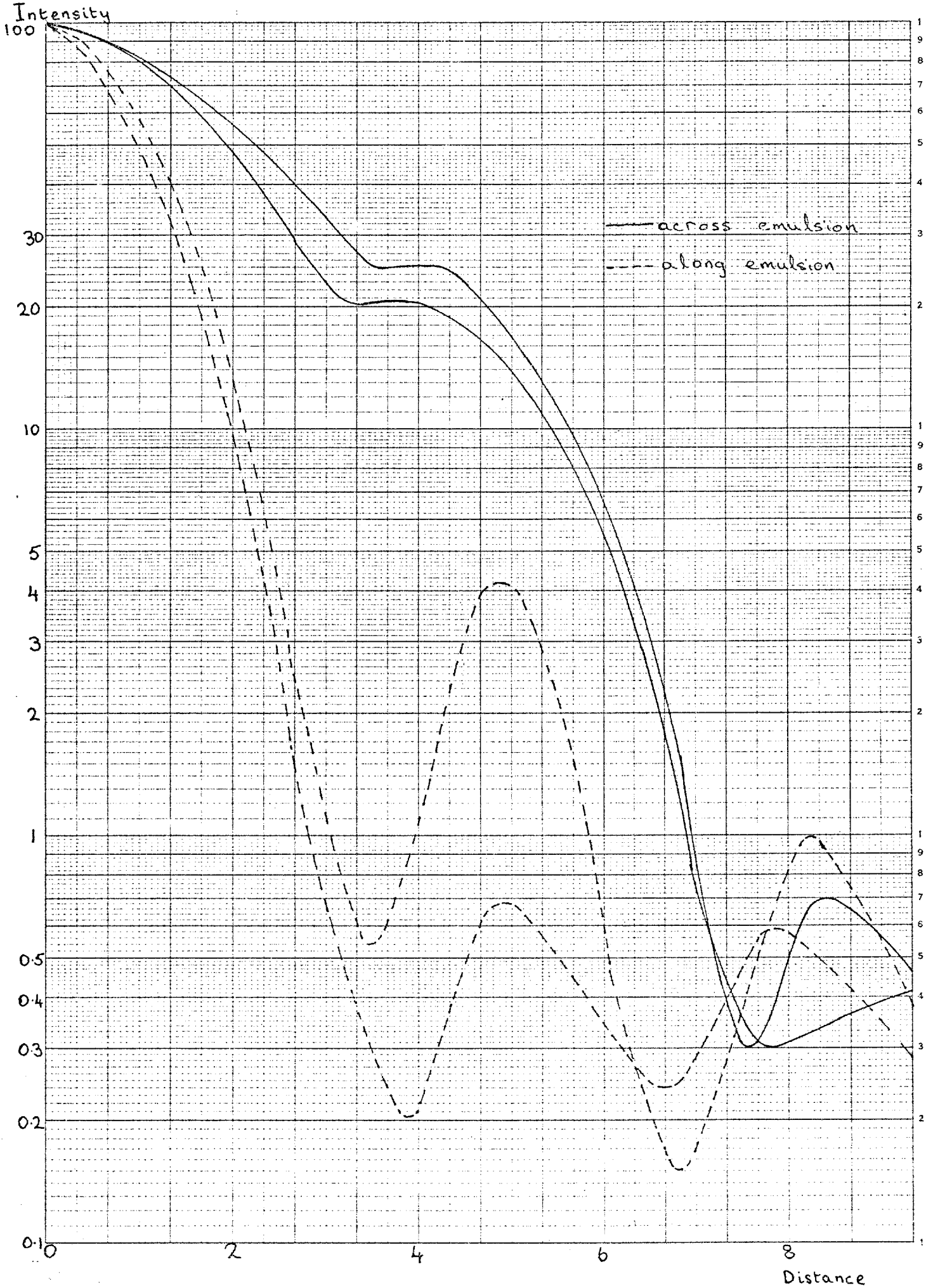


Fig 33 78 sq. mm. apt.

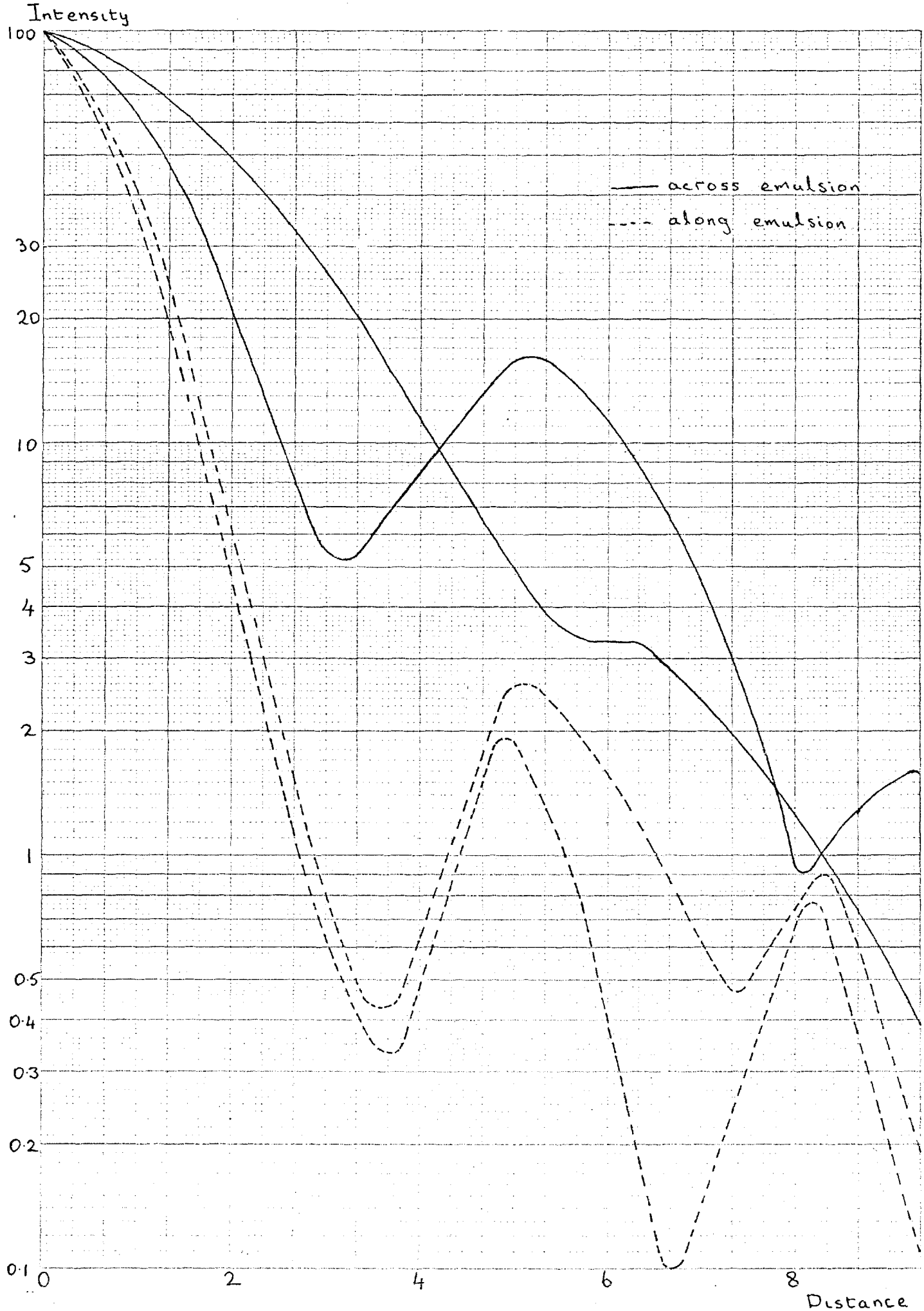


Fig 34 78 sq. mm. apt.

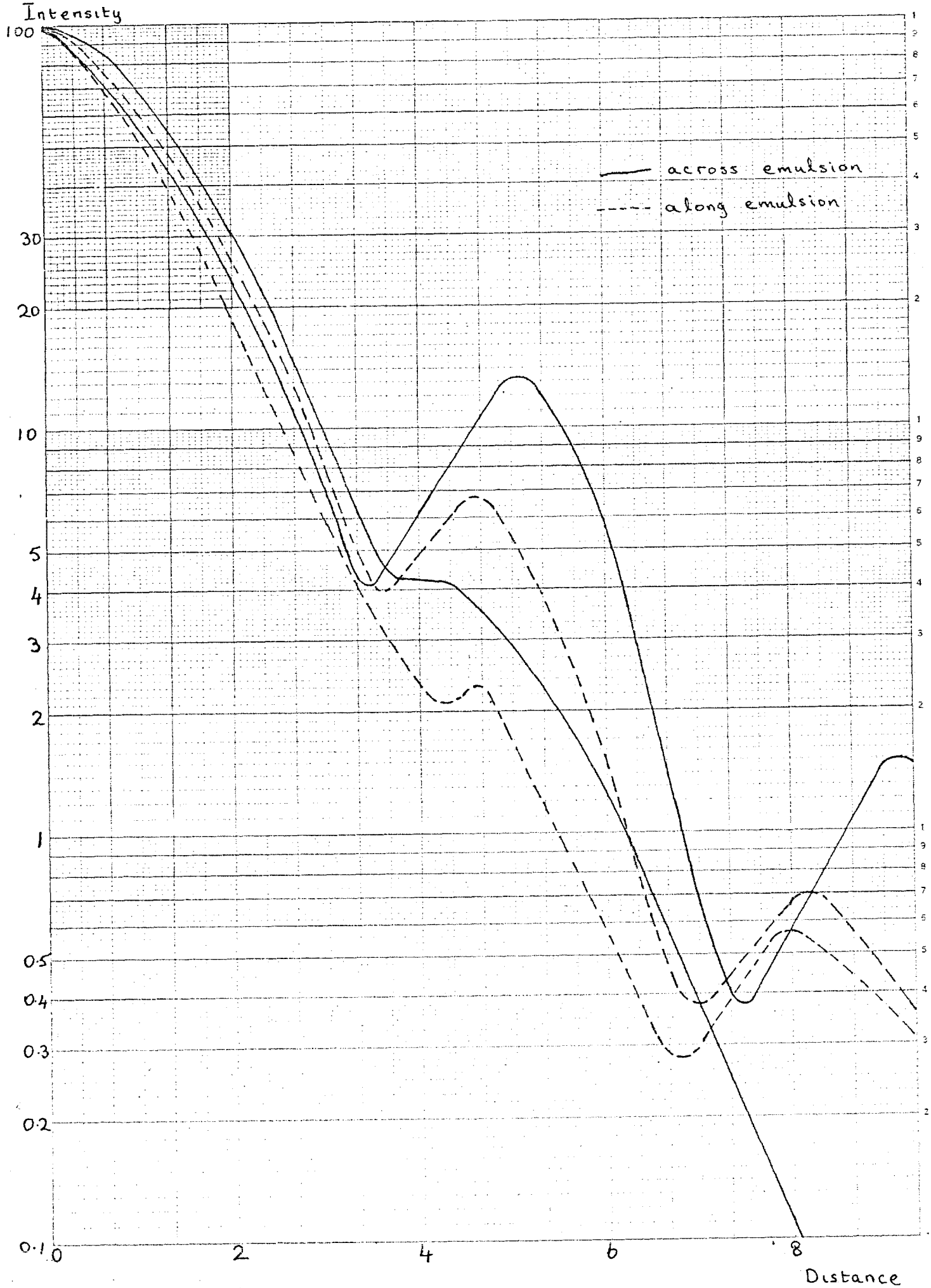


Fig 35 78 sq. mm. apt.

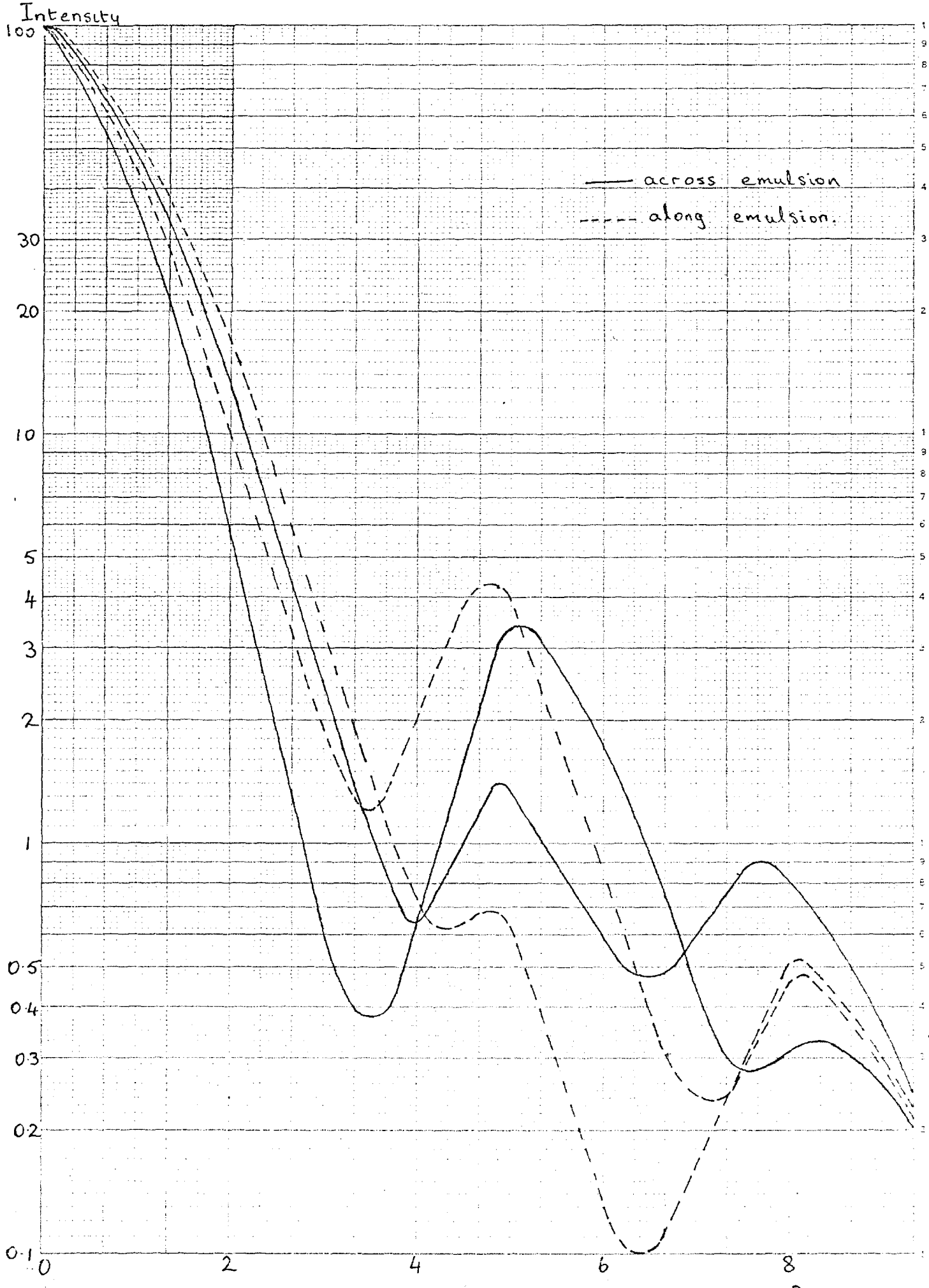


Fig 36 78 sq. mm. apt.

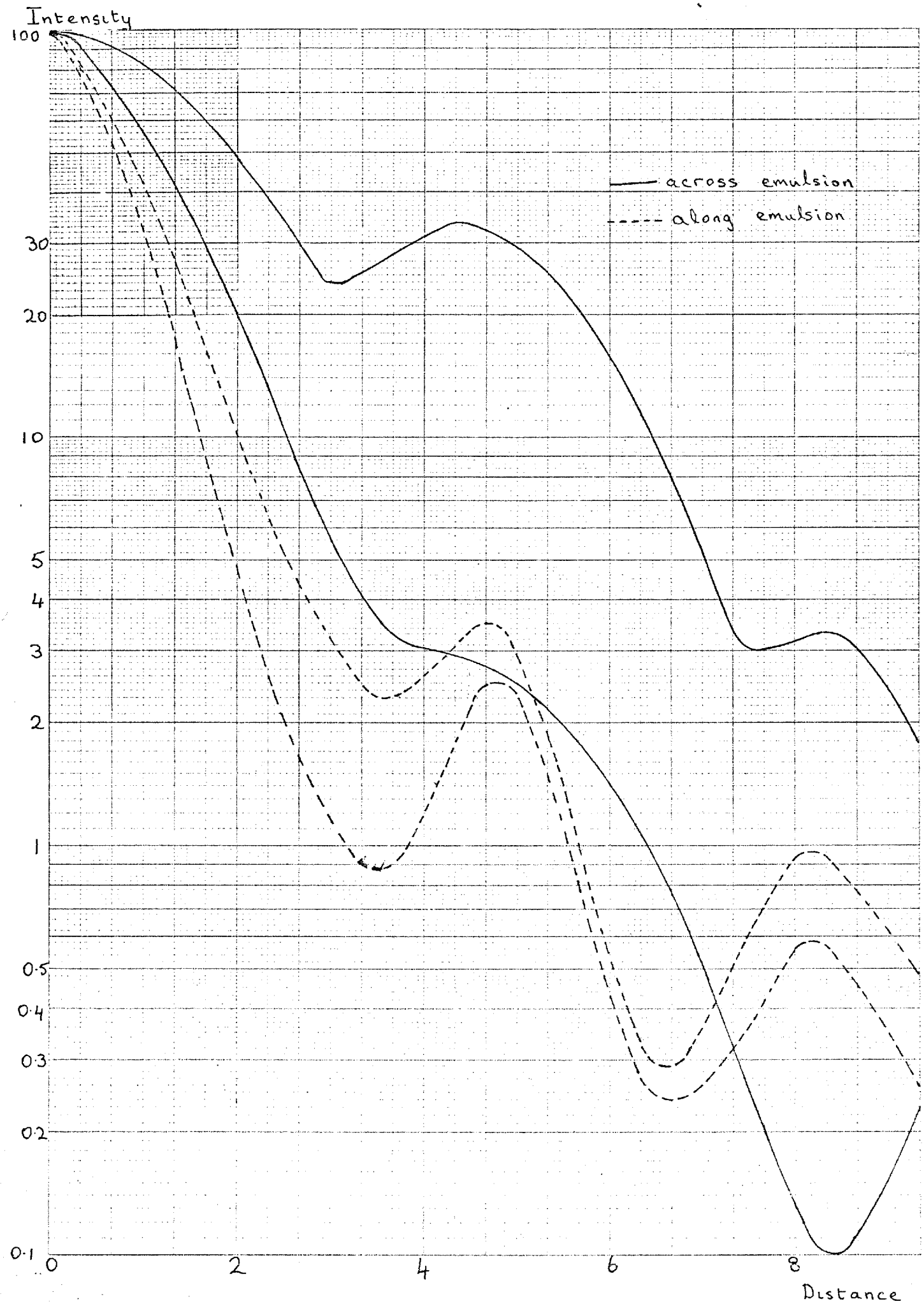


Fig 37 78 sq. mm. apt.

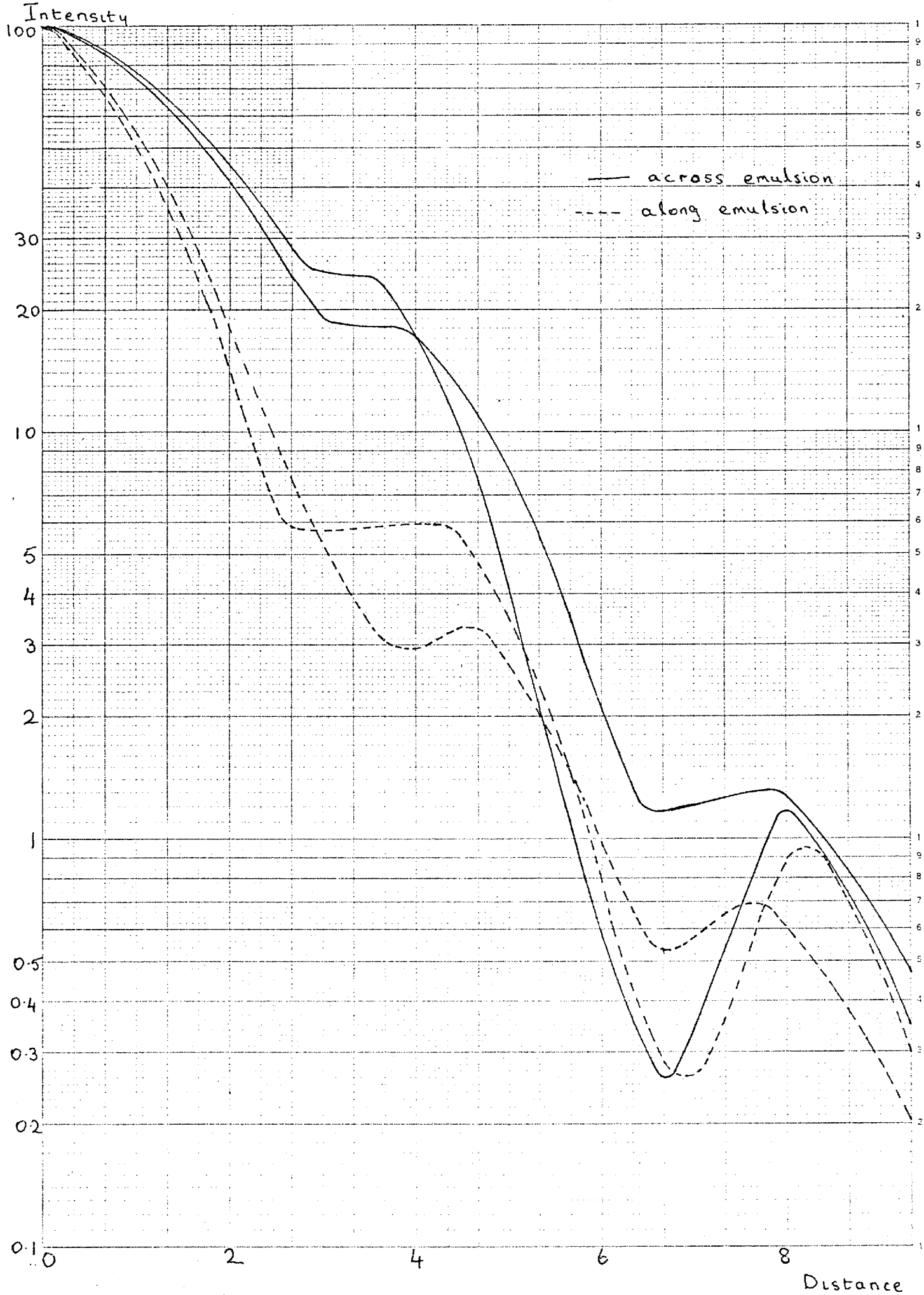


Fig 38 78 sq. mm. apt.

within the tolerance criterion, and another portion was just about acceptable. The rest had a variance ranging between $0.01\lambda^2$ and $0.02\lambda^2$. With this smaller area it is also expected that the number of scratches is greatly reduced.

(i) Fig. 40. The scan along the emulsion agreed well with that of the diffraction limited case. Across the emulsion scan showed one relatively high peak of about 4.5%. The value of the variance, $0.006\lambda^2$ agreed well with the pattern obtained.

(ii) Fig. 41. A variance of $0.02\lambda^2$ was reasonable to expect from the figure obtained. Three of the four first maxima were higher than the diffractionlimited case, the highest of these being at 16%.

(iii) Fig. 42. A symmetrical scan was obtained along and across the emulsion, with the maxima and minima in the correct positions. The scan across the emulsion had first maxima of nearly 9% as was indicated from a variance of $0.01\lambda^2$. Along the emulsion the corresponding peaks were at about 2%.

(iv) Fig. 43. This portion of base had a variance within the tolerance limit. The Airy disc obtained showed this by having most of the light energy within the central maximum.

(v) Fig. 44. Very similar to Fig. 42 but with slightly higher maxima. As expected the variance was also a little higher, $0.015\lambda^2$.

(vi) Fig. 45. Almost identical with the previous figure with the same value of variance.

From the results of the twenty scans it is seen that the random phase variations have random effects on the diffraction plane. The central maximum can be too broad or too narrow. Sometimes we have well defined maxima and minima and sometimes we just have a rather large spread of light with no real distinct first maxima or minima. The scans on the smallest size area were by far the most consistent.

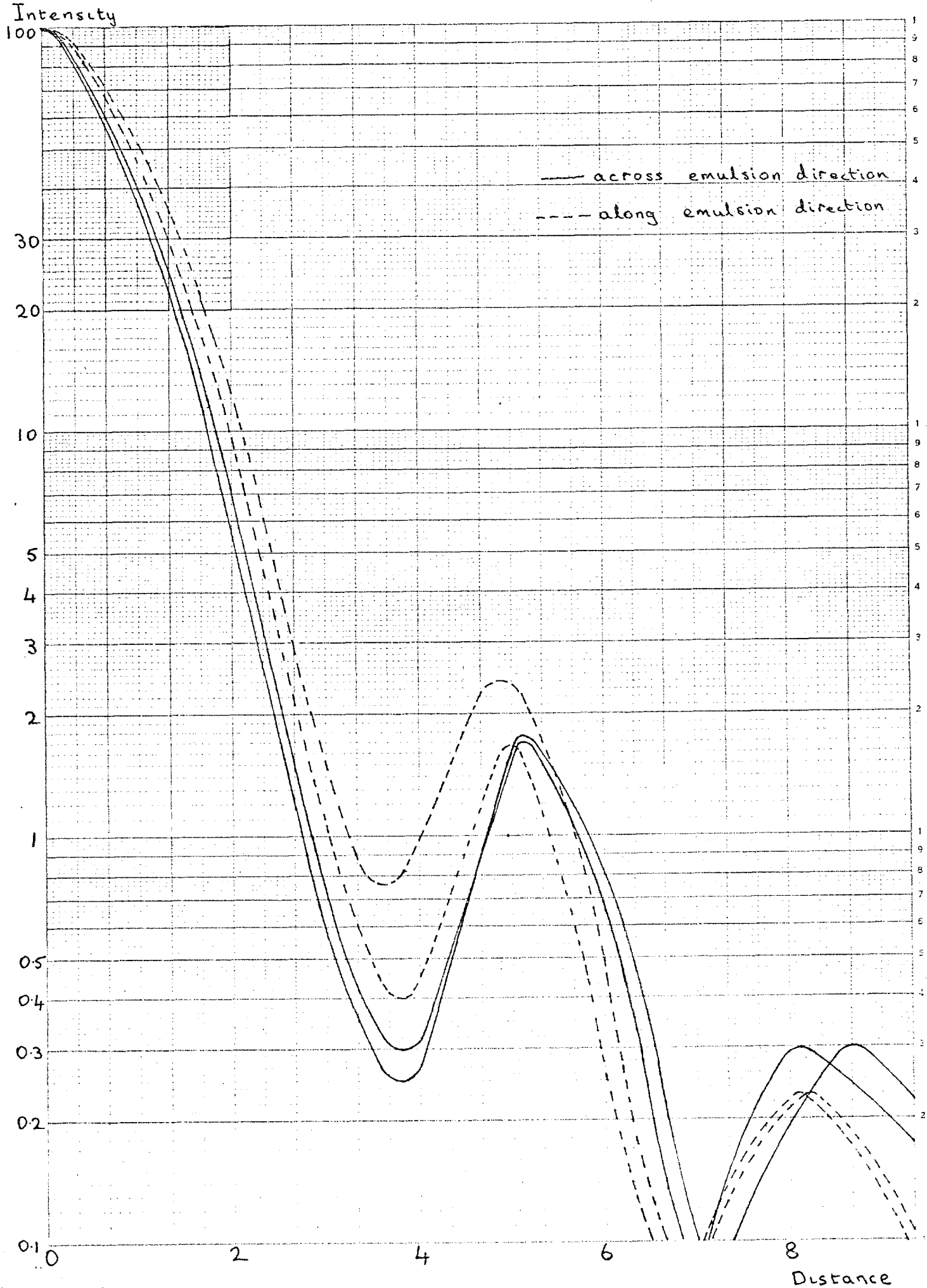


Fig 39 No Emulsion 38 sq. mm. apt.

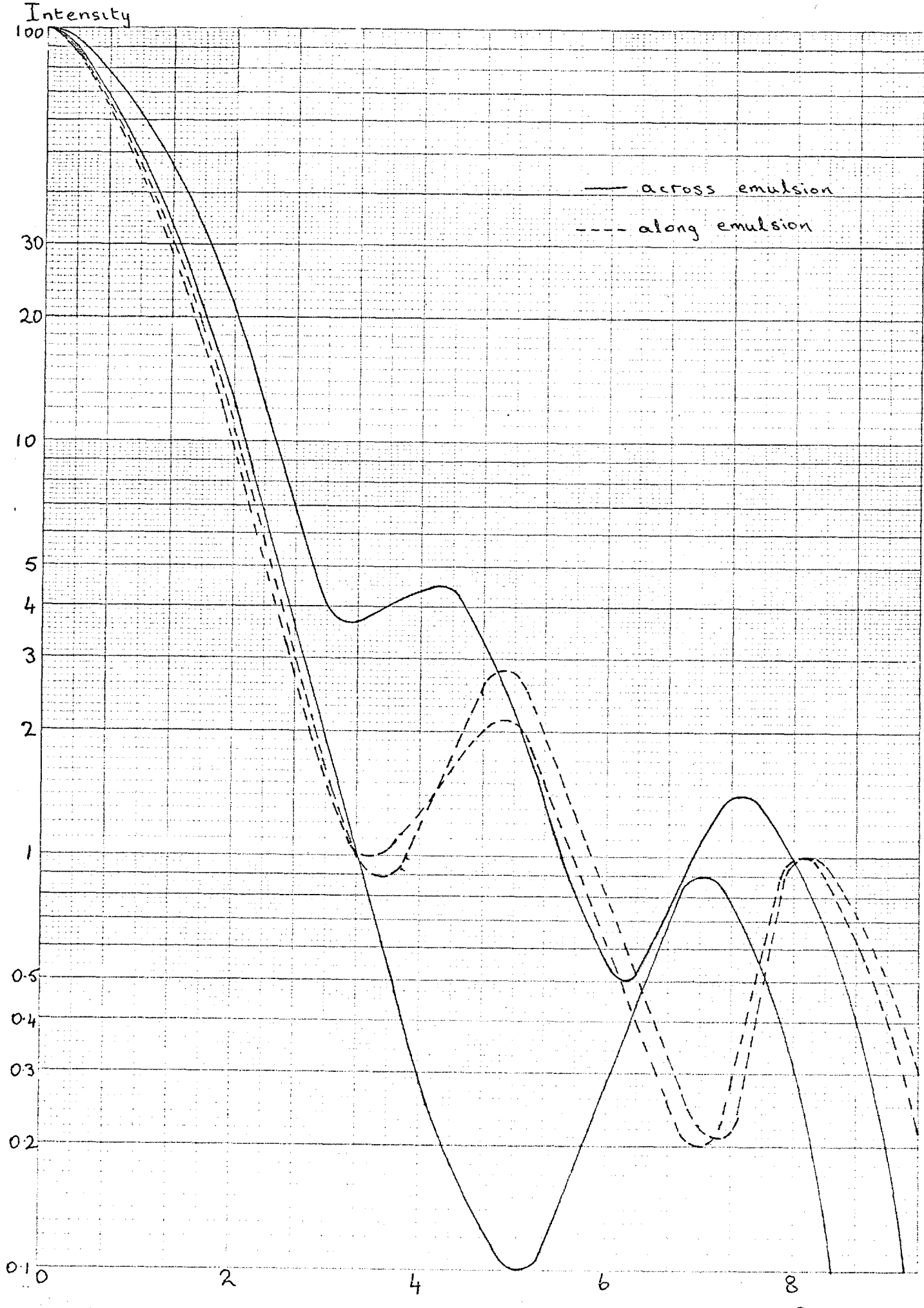


Fig 40 38 sq mm. apt.

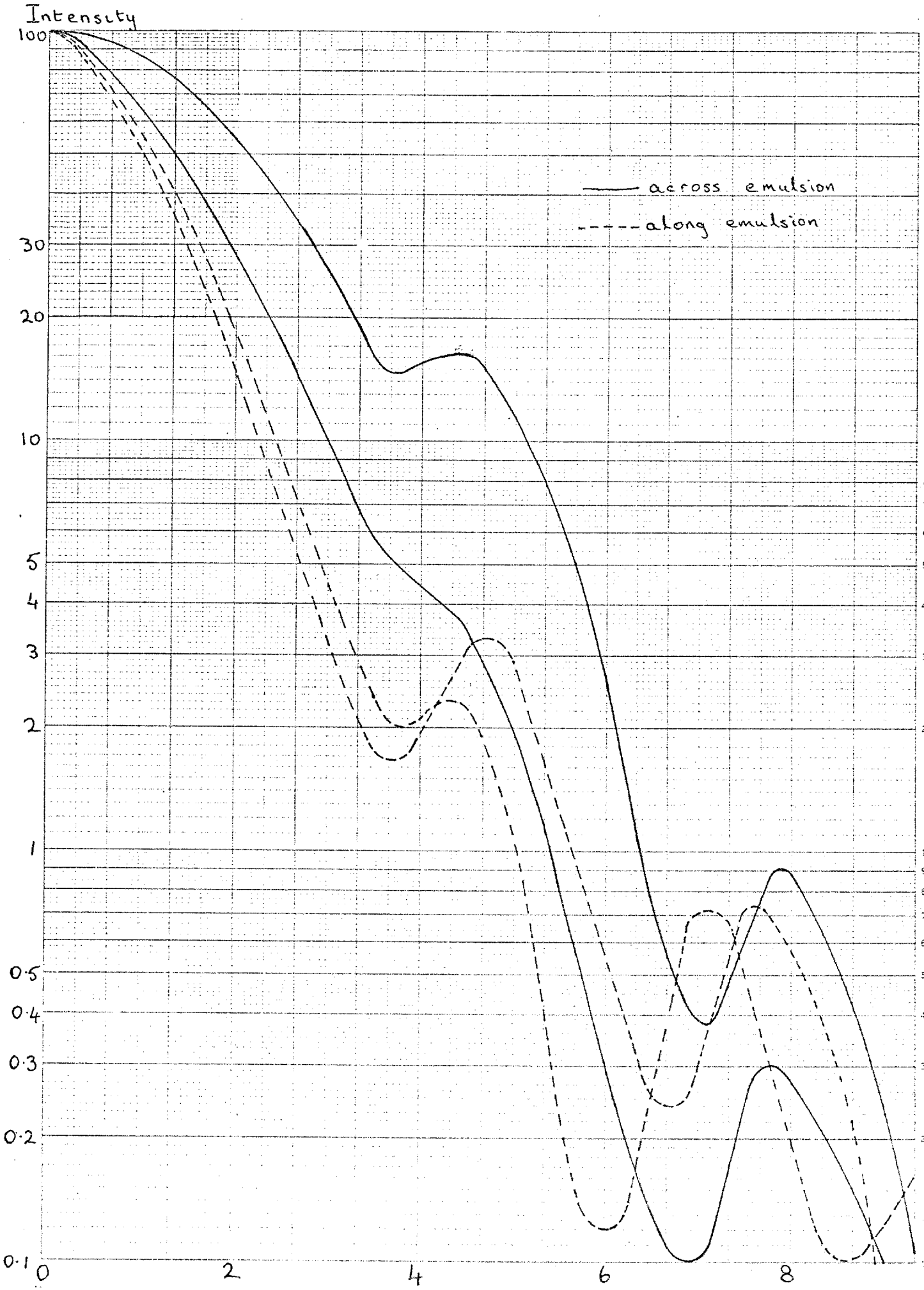


Fig 41 38 sq. mm. apt.

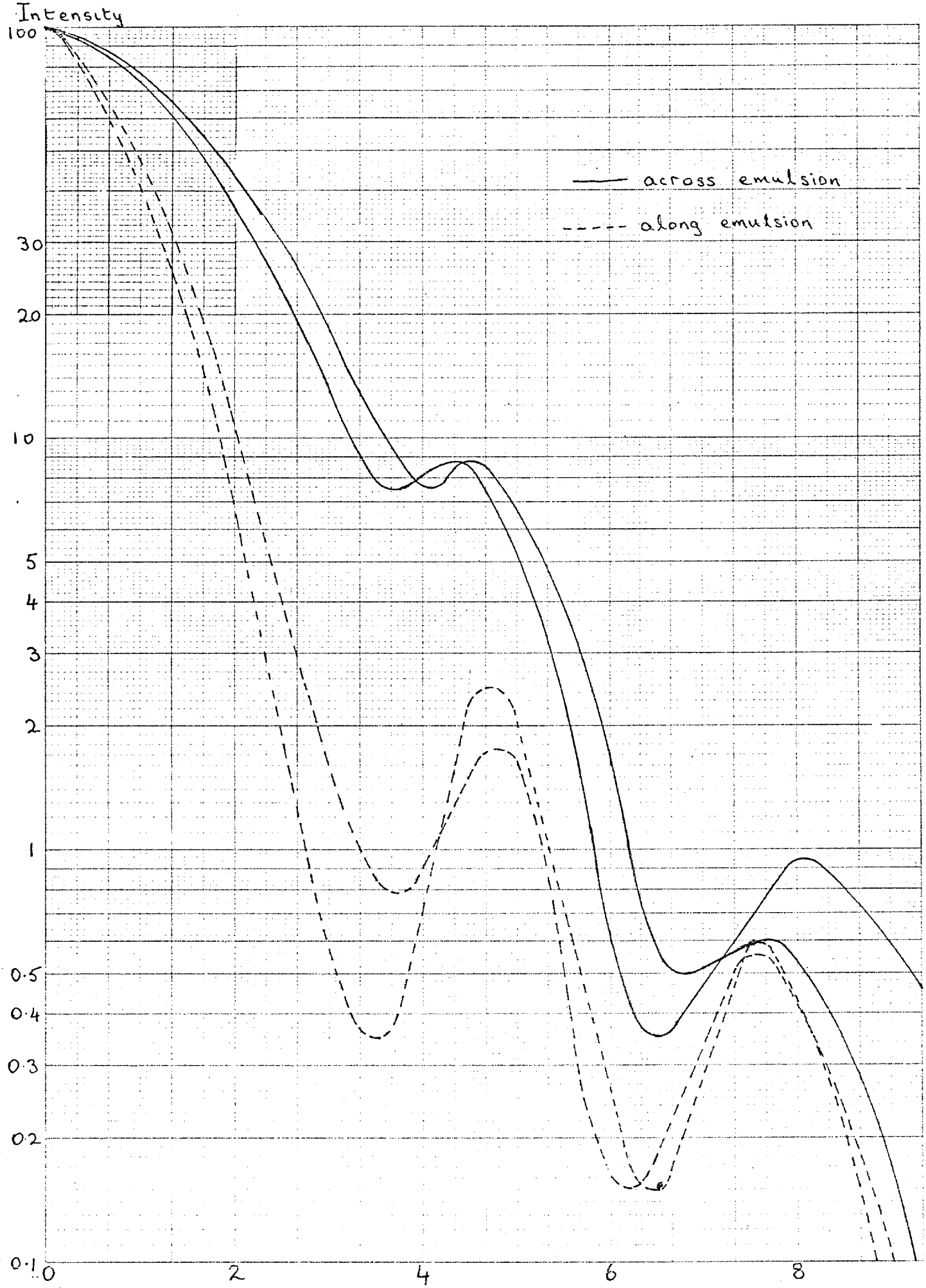


Fig 42 38 sq. mm apt.

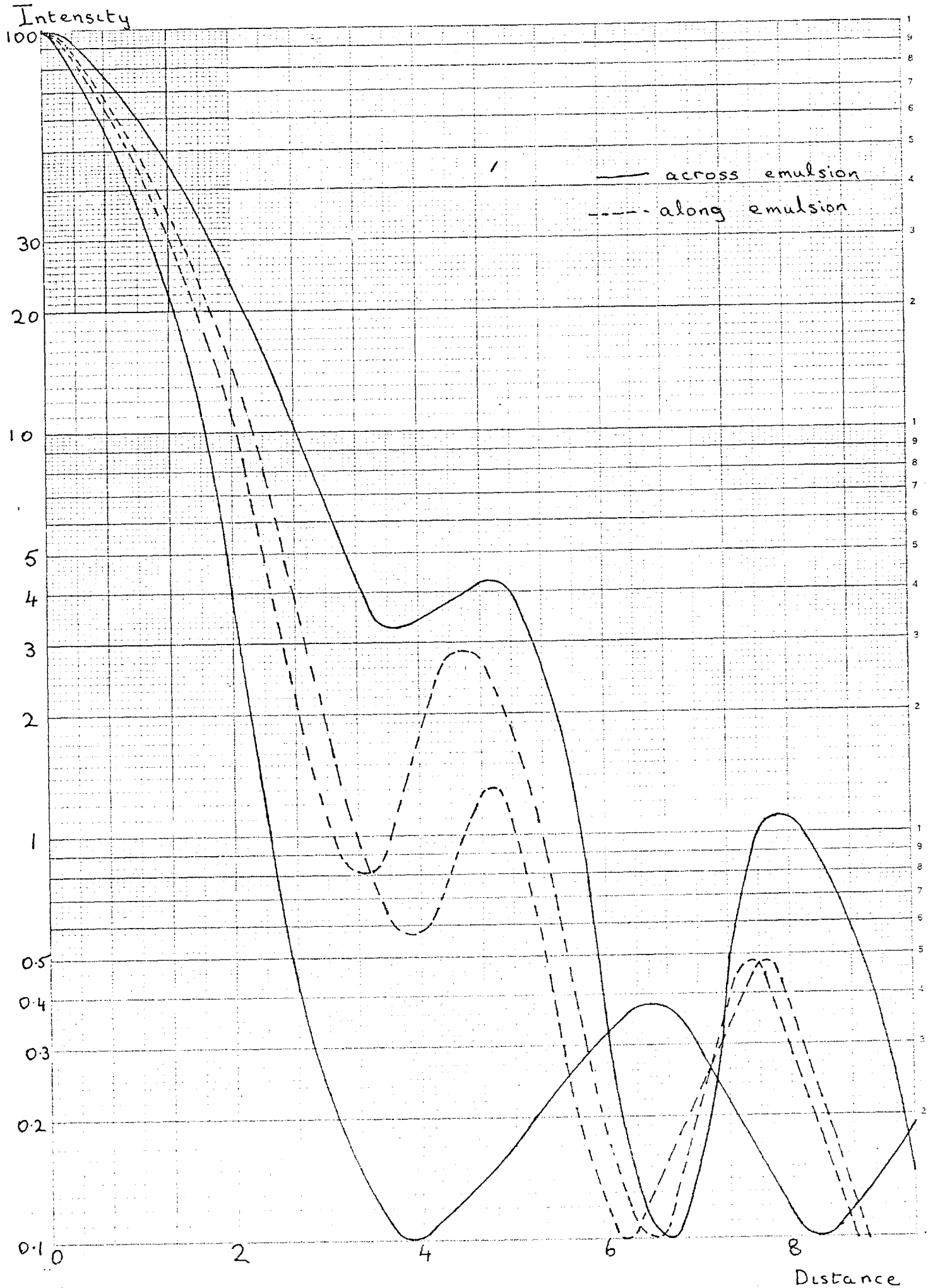


Fig 43 38 sq. mm. apt.

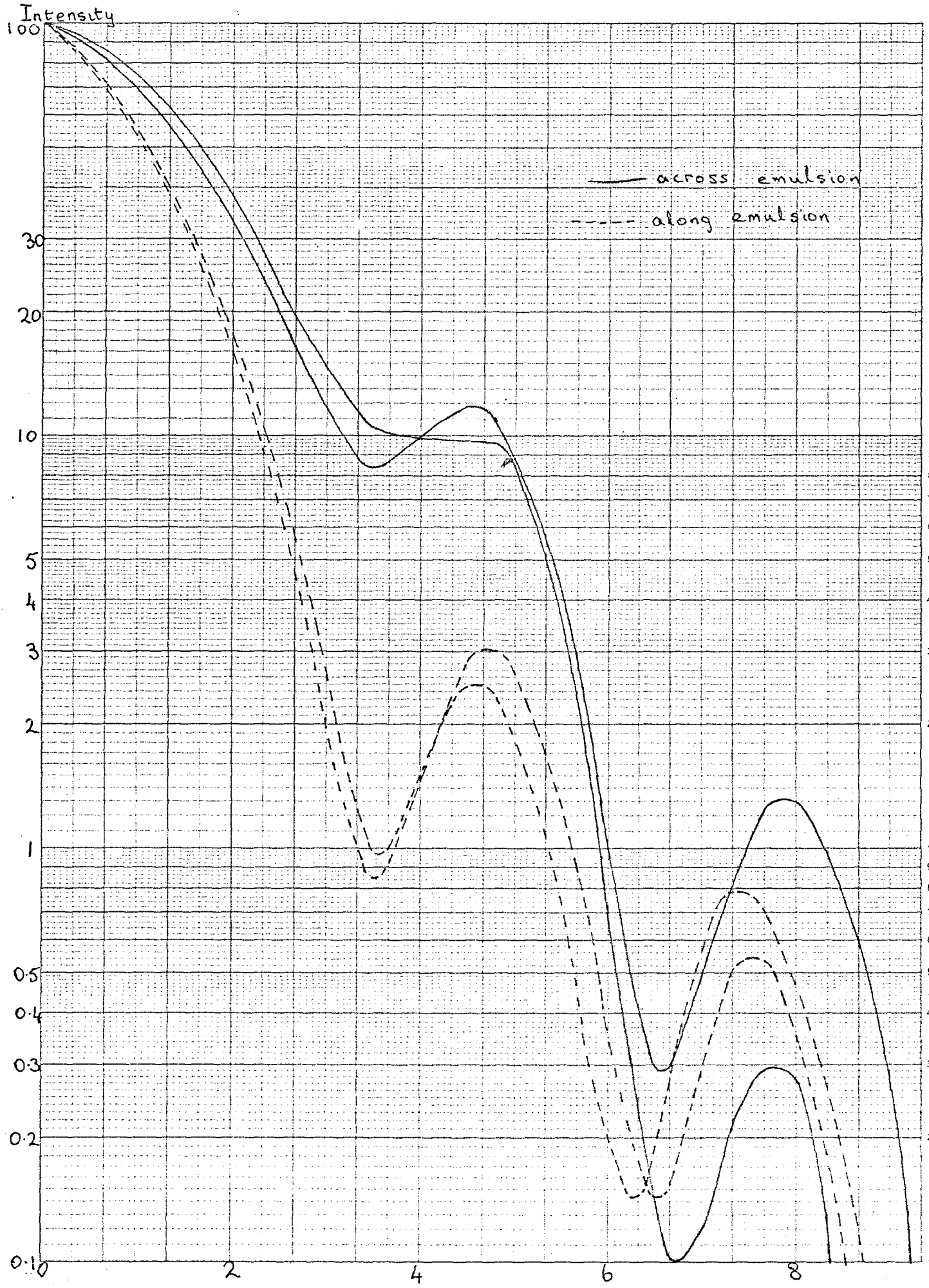


Fig 44 38 sq. mm. apt.

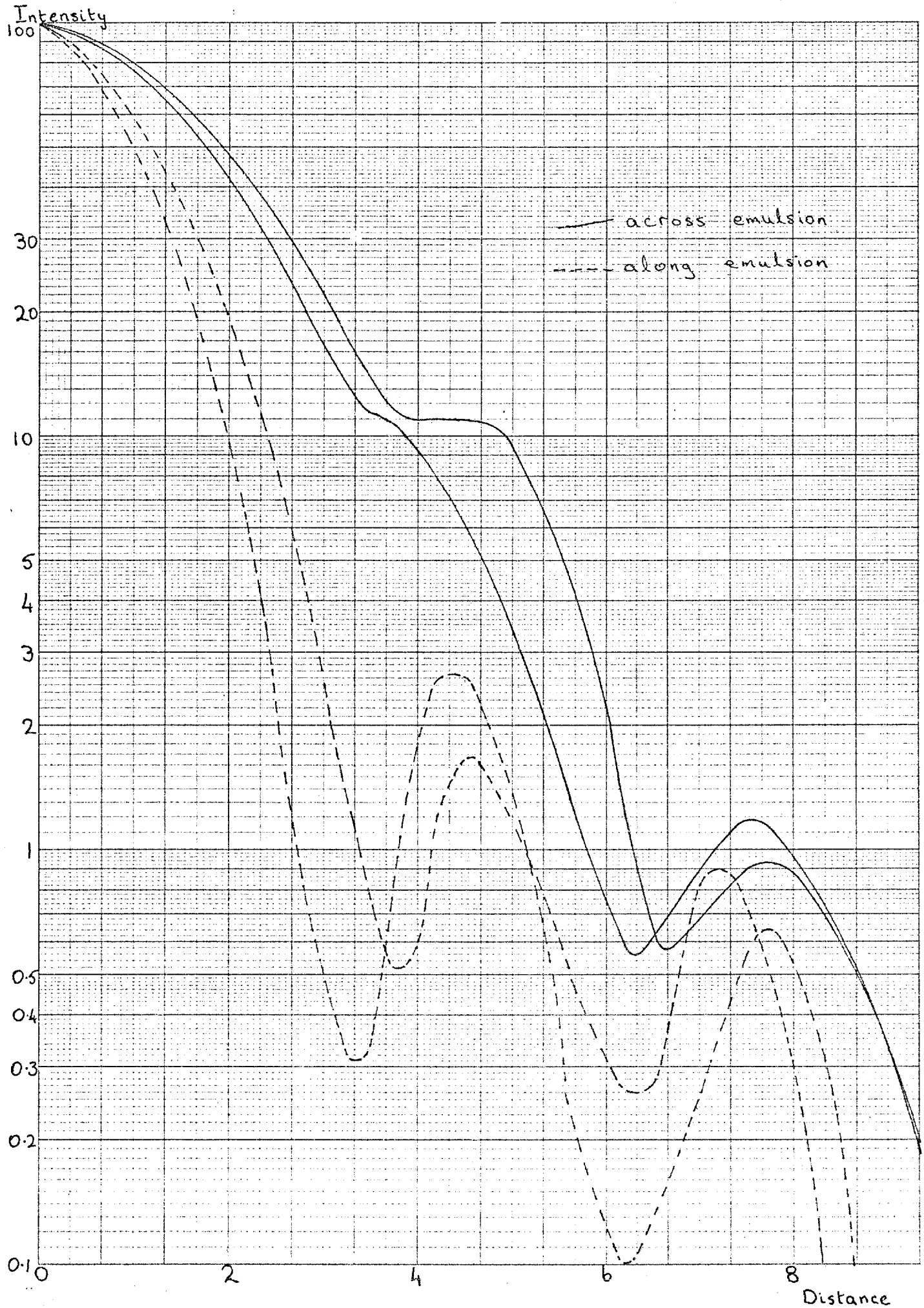


Fig 45 38 sq. mm. apt.

The positions of the maxima and minima occured in the correct places and it was easy to see how the val ue of the variance was related to the peaks of the first maxima. The value of these peaks increased with the variance. It was encouraging to see that the number of acceptable areas agreed with the expected, Fig. 13, despite the fact that we only had a limited number of samples.

V.7 Scatter

Thus far we have been considering the emulsion as having amplitude and phase properties. These characteristics are stored in the light sensitive grains of the emulsion which is held together in a suspension of gelatin. Because of the finite size of the grain structure the images formed will be degraded. The study of the granularity, signal to noise ratio and information sensitivity is well known and will not be considered here. The main reason for this is that our investigations are mainly concerned with the use of the emulsion in optical processing systems. The imaging properties of the emulsion are of course relevant and these will be considered in the next section on the modulation transfer function.

Owing to the granular structure of the emulsion when the emulsion is developed to give a constant macroscopic transmittance we will also have a microscopic transmittance. The resultant of this effect is that a portion of the light is diffracted or scattered by the emulsion into an angular spectrum corresponding to the spatial frequency of the fluctuations. When viewed in the diffraction plane this light has an intensity which represnts the Wiener spectrum[39] of the microscopic transmittance. For coherent illumination the intensity in the diffraction plane is given by the square of the Fourier transform of the amplitude transmittance variations and not the deviation from the

mean intensity transmittance. If there are any phase variations associated with the microscopic amplitude variations then these could also scatter the light. The resulting effect is that this light interferes with the signal recorded on the film and can thus be regarded as noise. This noise which is due to granularity is a function of both density and development conditions. For our present case we kept the development conditions the same for all cases and investigated the effect of the density.

So far we have only mentioned the scattering due to the silver grains. At low angles the effect of the gelatin should also be considered.[6] For our case the thickness of the gelatin layer is small and hence it does not have aslarge an effect as most other emulsions. Even so, it must be realised that the resultant effect is due to both the silver grains and gelatin.

The apparatus used for these experiments was the same as that used in the previous section. Various uniform densities of emulsion, ranging from 0.4 to 1.5, were placed in the input plane of the processor. The resultant Airy disc pattern was scanned with a five micron pinhole. In order to avoid confusion with the scatter produced by scratches, only scans along the length of the emulsion were considered. As before the scans were normalised with those obtained without any emulsion present. In order to have a reasonable light level, especially when using high density emulsions a rather large aperture, 1.5 cm diameter, was used to illuminate the emulsion. The scans were taken over three Airy rings, the amount of scatter is given as a percentage of the central maximum. The results of the variation of scatter with density are shown in Fig. 46. From this diagram it is seen that the effect of the emulsion is to produce a background scatter, the amount of this scatter increasing with density. For densities

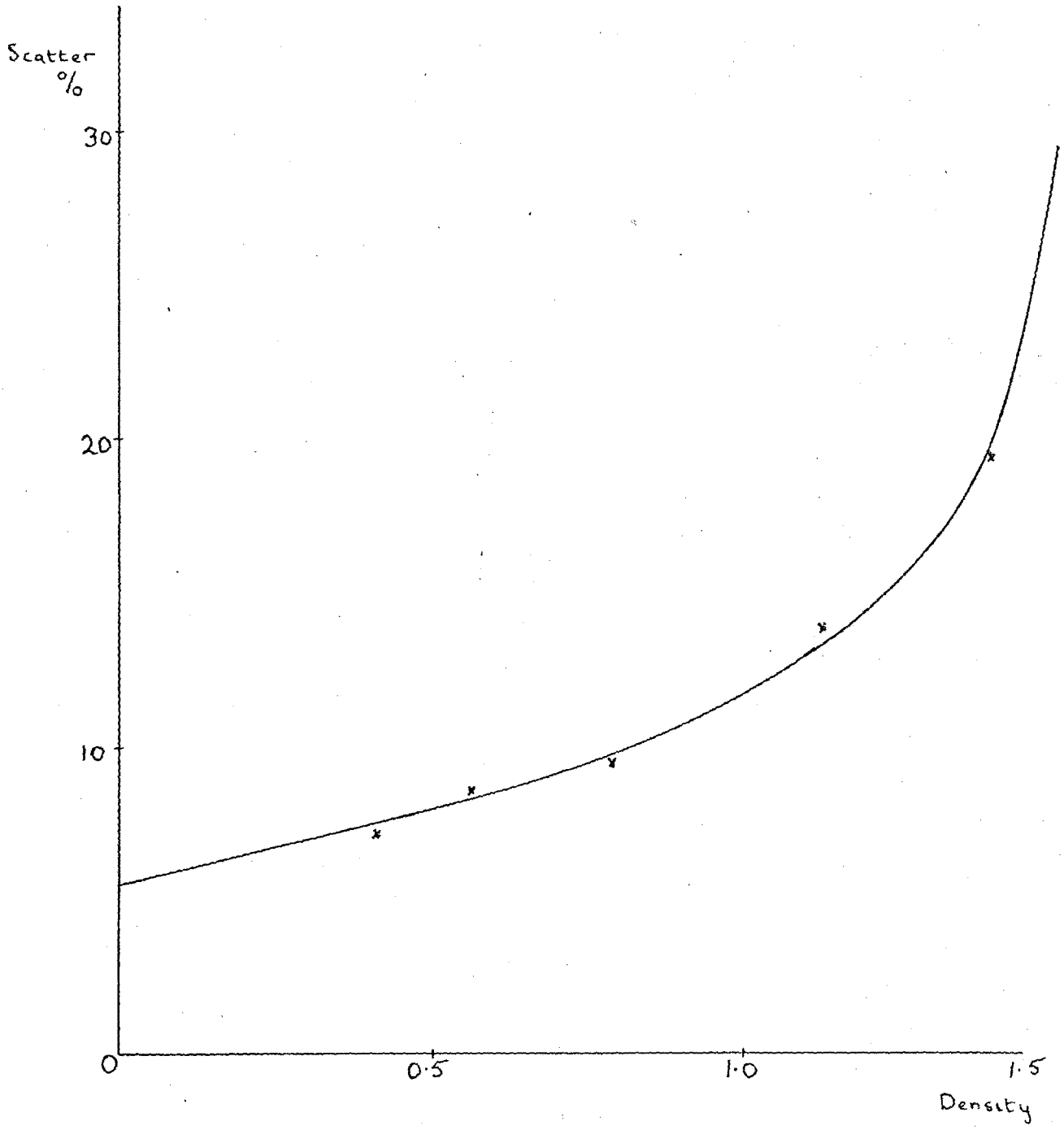


Fig 46 Scatter

below unity the scatter increases slowly and linearly with increasing density. After about this point however, the amount of background scatter seems to increase much more rapidly with increasing density. At a density of 1.5 the amount of scatter was so high that no secondary peaks were observed in the Airy disc pattern. If we extrapolate Fig. 46 back to find the amount of scatter at zero density we can assume that this is due to the gelatin and base. This value is seen to be of the order of 5%. From the scans of the base alone it is seen that virtually all of this is due to the gelatin. Therefore up to density of 0.8 the majority of the scatter is from the gelatin. At this density we have an equal contribution from the silver grains and gelatin. After this density the silver grains give rise to the greater part of the scatter.

As has already been mentioned the position at which the scatter was measured corresponded to the third Airy ring, which for our case was at a distance of 40 microns from the central peak. For some of the scans the distance of scan was extended to 100 microns. From these further scans it was seen that the scatter did not drop noticeably for the low density emulsion but was more noticeable for the higher density emulsion. This seems to indicate that the extent of the gelatin scatter does not change over this range of distance, but that due to the silver grains appears to fall.

A further fact which was seen from the scans was that the amount of scatter was symmetrical about the central maximum. This is expected since it is assumed that the amplitude transmission fluctuations are random and so they would not give a preferred scatter direction.

V.8 Modulation Transfer Function

As a conclusion to examining some of the optical properties of

the Ilford RX emulsion, it was decided to take a brief look at the frequency response of the emulsion. In Section 2.6 we discussed the modulation transfer function and mentioned the condition for linearity. Two important factors which must be taken into consideration when using Fourier transform techniques for photographic emulsions are (a) the signal amplification introduced by the emulsion, and (b) the linearity of the emulsion. The question of signal amplification is easily dealt with since it is measured by the gradient of the characteristic D-logE curve, i.e. the gamma. The second factor, that of linearity, is far more difficult to deal with. Linearity for the photographic emulsion implies a linear relationship between density and the logarithm of exposure. This is not however, equivalent to the Fourier transform definition of linearity, since for the photographic case a sinusoidal exposure does not necessarily produce a sinusoidal image when operating according to its linearity condition. For the photographic case it would perhaps be more meaningful to speak in terms of linearity between transmission and exposure, since this would then also be meaningful in Fourier transform applications. Linear recording of sine wave objects was mentioned by Selwyn as early as 1948.[68] The result of this non-linearity is to introduce harmonic distortion into the recorded signals. Little[59] shows that the harmonic distortion of sinusoidal objects can be rendered insignificant by processing to a positive under certain conditions by exposure and development. Due to the time factor involved in processing to a positive, there would be little point in adopting this procedure for rapid access emulsions. It is however reasonable to assume that over a limited range the amount of harmonic distortion introduced in the negative process can also be reduced to a negligible

amount.

There are several methods whereby the modulation transfer function can be measured for photographic emulsions. We can use either a non-periodic object such as a slit or an edge of we could use a periodic object such as a sine wave or a square wave. In an attempt to reduce the amount of harmonic distortion and minimise the adjacency effects, it was decided to use the sine wave response as a means of measuring the transfer function. Low contrast sinusoidal two beam interference fringes were exposed onto the emulsion at a bias intensity transmittance corresponding to the centre of the D-logE linear region. These were then developed in Phentrace developer at 20^{o}C with constant agitation, to minimise the adjacency effect. The developing time was adjusted to give a gamma of one. The resulting density fluctuations were then scanned using a Joyce Loebl microdensitometer. From these traces the contrast was evaluated for each of the frequencies. The recorded frequencies ranged from 0.5 c/mm up to nearly 30 c/mm. It is reasonable to assume that at a frequency of 0.5 c/mm the response of the emulsion is virtually 100% and so it was assumed that the object contrast was the same as that given by the 0.5 c/mm value. This value was therefore used to normalise the sine wave response for the other frequencies. The results of this were used to plot the modulation transfer function curve of Fig. 47.

It is realised that the transfer properties of the optical systems involved in recording and scanning the signal have been ignored. This is because their effect is small for the low frequencies considered and the curve obtained is a good representation of the emulsion performance. From this curve it is seen that the emulsion performance is reasonably good and of the same order as other emulsions of a similar speed. This curve has been obtained for the best development conditions and it is

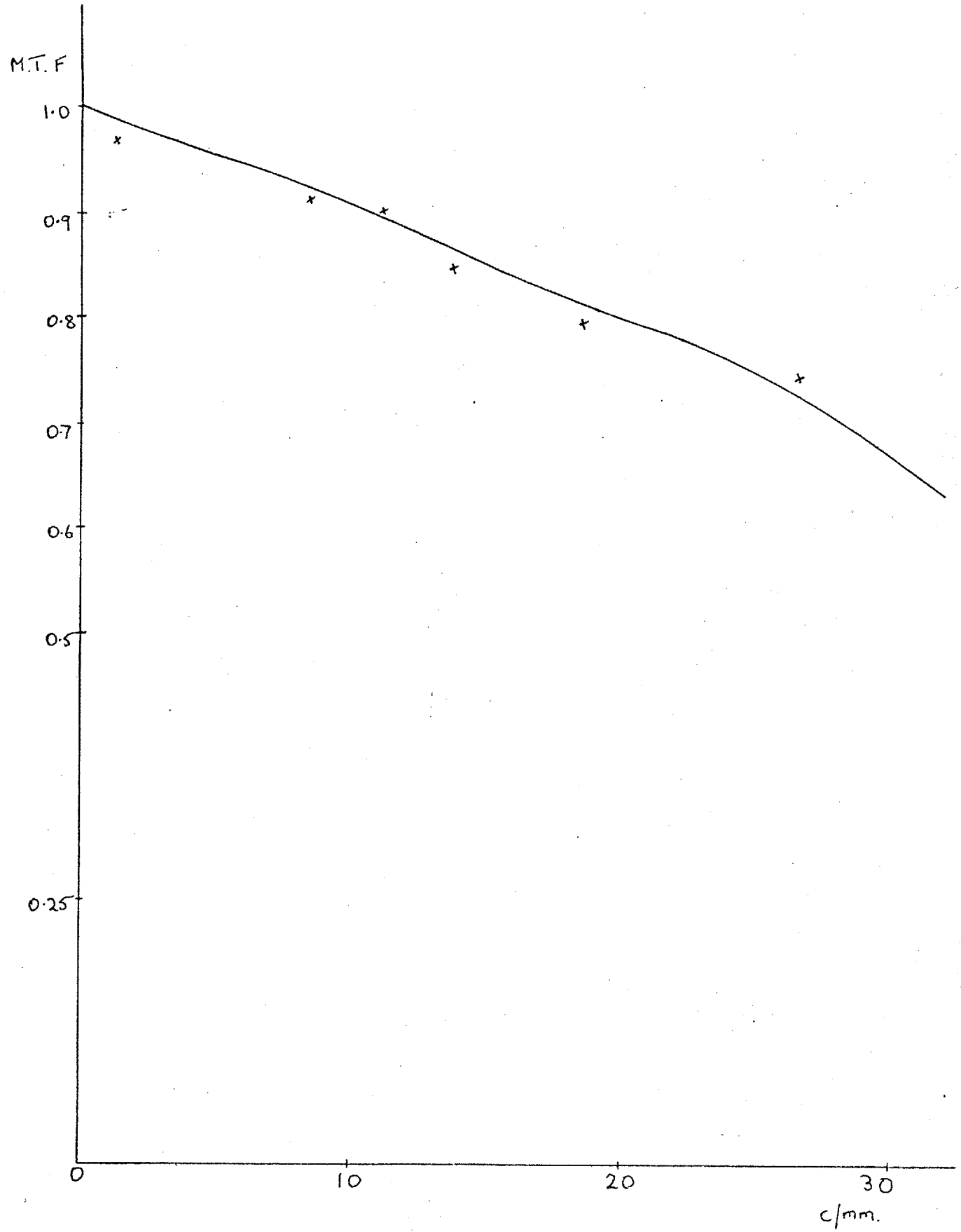


Fig 47 M.T.F.

unlikely that such a response can be expected for rapid processing conditions. Unfortunately, with the conventional processing equipment available it was not possible to perform the same experiments with any degree of accuracy under rapid processing conditions (less than 6 seconds).

CHAPTER VI - LIQUID GATES

In the previous chapter we demonstrated the presence of phase variations in photographic emulsions. In the case of the random emulsion phase variation, their effect was to degrade the image, and hence their presence is unwanted. The other type of phase variation, introduced by the relief image, has the effect in some instances of increasing the diffraction efficiency of the emulsion. Thus this type of phase variation can improve the emulsion performance. On the other hand, for certain particular relief heights this phase image can completely change the expected Fourier transform as mentioned in 2.7. A liquid gate is a device used to surround the film with a refractive index matching liquid. If the phase variations are the result of film thickness changes, then the liquid gate could be used to reduce them to a negligible amount (they could not be completely eliminated for the present film since the base and emulsion have a different refractive index). If the variations are due to refractive index changes then the degree to which the phase variations can be eliminated would depend upon the extent of the refractive index changes. The object of this brief investigation was to see whether the phase variations could be eliminated by using a liquid gate, and the extent of the refractive index changes.

A further important advantage of a liquid gate is that it would tend to reduce the amount of noise due to scratches on the film. The disadvantages arise in the difficulties in handling of the liquid and film when using such a liquid gate.

VI.1 Design of the Liquid Gate

The basic design of a liquid gate is very simple. All that is required is a suitable liquid and two optical flats. The optical flats, 3 cm diameter, 0·7 cm thick and flat to within one tenth of a wavelength in transmission, were cemented into two aluminium blocks, as shown in Fig. 1. The two blocks were bolted together with a rubber seal between them. The liquid was poured into the cavity between the two blocks. The film is introduced between the rubber seal and positioned between the optical windows.

The final stage in the design of the liquid gate is the choice of the liquid.[36] Of prime importance is its refractive index. We know that the refractive index of the base and gelatin are different. However we also know that most of the random phase variations are due to the base. Hence, in order to eliminate these phase variations we require a liquid whose refractive index closely matches the base. Our tolerance criterion is defined by the Strehl ratio which is equivalent to saying that the optical path difference in transmission is less than one quarter of wavelength, of the light, over the entire aperture. Now if we let n_1 be the refractive index of the film and n_2 that of the surrounding medium, then the optical path difference due to film thickness changes is given by

$$\text{O.P.D.} = (n_1 - n_2)t \qquad (6.1)$$

If we now allow the maximum permissible optical path difference then 6.1 becomes

$$\frac{\lambda}{4} = (n_2 - n_1)t \qquad (6.2)$$

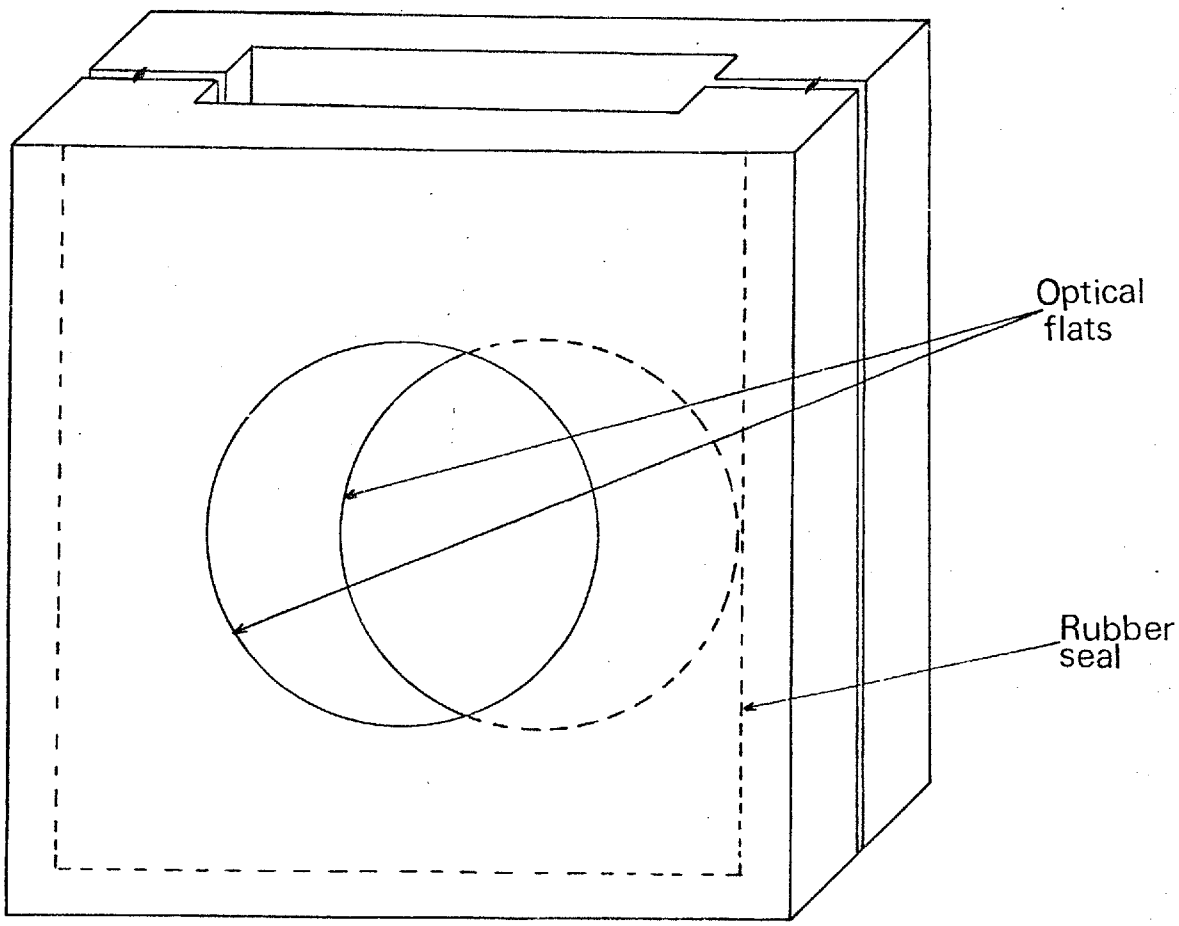


Fig 1 Liquid Gate

where λ is the wavelength of the light used. From 6.2 we see that the closer we match the refractive index the more thickness variation t we can tolerate and still be within our criterion limit. Fig. 2 shows the permissible thickness variation as a function of matching liquid refractive index when using the red laser line. From this figure we see that if the film was immersed in water then we could tolerate about one micron of thickness variation. As the matching liquid approaches the film refractive index then the amount of thickness variation that we can tolerate increases rapidly. For a refractive index match of 0·01 we can tolerate nearly 16 microns of thickness variation. If we were to match the refractive index of the base exactly, then we could have any amount of base thickness changes. Matching the base gives a missmatch with respect to the gelatin of 0·04 which in turn means that we have less than about 4 microns of gelatin thickness changes to be within our tolerance limit. This thickness variation is greater than the actual thickness of the gelatin. Thus we see that for our case it is sufficient to match the liquid with the base to produce an optically flat system.

The second requirement of the liquid is dependent on its chemical properties. It must not in any sense react with the film thus changing its composition. It is also advisable for obvious reasons that the liquid should in no way be poisonous or harmful when taken externally. It should initially be colourless and during the period of its use it should not be discoloured. It would also be advantageous if it had a slow rate of evaporation.

The final consideration regarding the liquid is peculiar to the type of mechanical design used. The film is inserted in between the rubber seals. In order to examine a different part of the

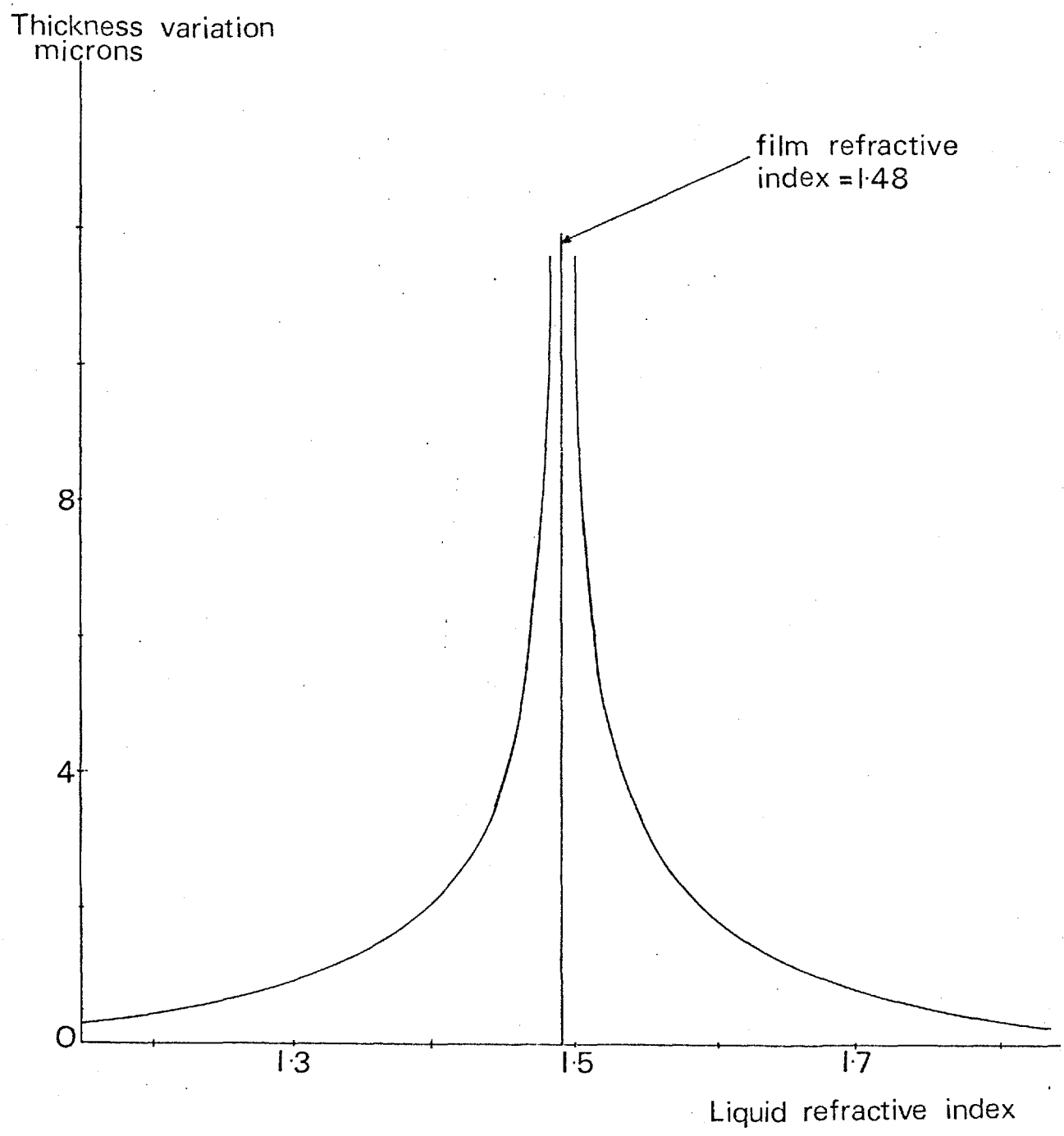


Fig 2 Refractive Index Matching

film it is simply pulled across the optical window. So that we can easily pull the film through the rubber seals it would be advantageous to have the seals as loose as possible with the liquid being a good lubricant. To avoid the liquid leaking away in this design, the viscosity of liquid should be relatively high. Its viscosity should be sufficiently high to minimise leaking but also low enough to fill in quickly any scratches on the film.

For the above reasons the type of liquid chosen was one of the range of silicone fluids. The properties of this fluid, a phenyl methyl silicone fluid, MS 550, seem to be ideal for our application. Firstly, it has a refractive index of about 1·49 which is good enough for our investigation. Since it is extremely stable under extreme temperature conditions and its temperature coefficient of refractive index is small and so its index can be considered constant over the possible range of temperatures used. If necessary a different refractive index can be obtained by mixing with another similar silicone fluid of a different refractive index. Chemically the fluid is extremely stable and inert; it may be heated in air at 250° C for well over 1,000 hours without gelling or decomposition. It will not attack the emulsion or base, it is non-corrosive and non-toxic and hence can be used in absolute safety. At room temperature the fluid has a viscosity of 125 cS, this was found to be sufficient to reduce the leakage to a negligible amount and also to fill in any scratches. It is also a good lubricant. Just as its refractive index-temperature coefficient can be ignored over the possible temperature range, so can its viscosity temperature coefficient. A further useful property is that its light transmission is extremely good over the whole of the visible spectrum.

Hence we see that this type of fluid is ideally suited for use in liquid gates. A full list of its properties can be found in MS Silicone Fluids publications G.1 and G.11.

VI.2 Effect of Liquid Gate

The main objective in using a liquid gate is to eliminate the phase variations present in the emulsion. If we can do this successfully then we will be able to predict the performance of the emulsion, in coherent light, with far more accuracy. The second advantage is that we will be able to use larger areas of the emulsion and so increase the amount of recorded information.

In order to investigate the effect of the liquid gate on the random phase variations, the Mach-Zehnder interferometer was used in conjunction with the coherent processor, as described in the previous chapter. The area of the emulsion being examined was limited by a 1 cm diameter aperture. This size was chosen since the rest of the optical system was in effect diffraction limited (Fig. 3), and yet the emulsion performance, due to phase variations and scratches, can be quite bad. As before, tilt fringe interferograms were taken of the emulsion and the corresponding Airy disc pattern was scanned in two directions. From the interferograms it was seen that all the areas examined had a variance of wavefront well within the Strehl limit. This was confirmed by the scans of the Airy disc; Figs. 4 and 5 are two typical curves obtained. When these figures are compared to Fig. 3, the scan obtained without the emulsion, it can be seen that within the experimental limitations the emulsion has no effect on the distribution of light intensity in the diffraction plane. It was also seen that the effect of scratches, which was so

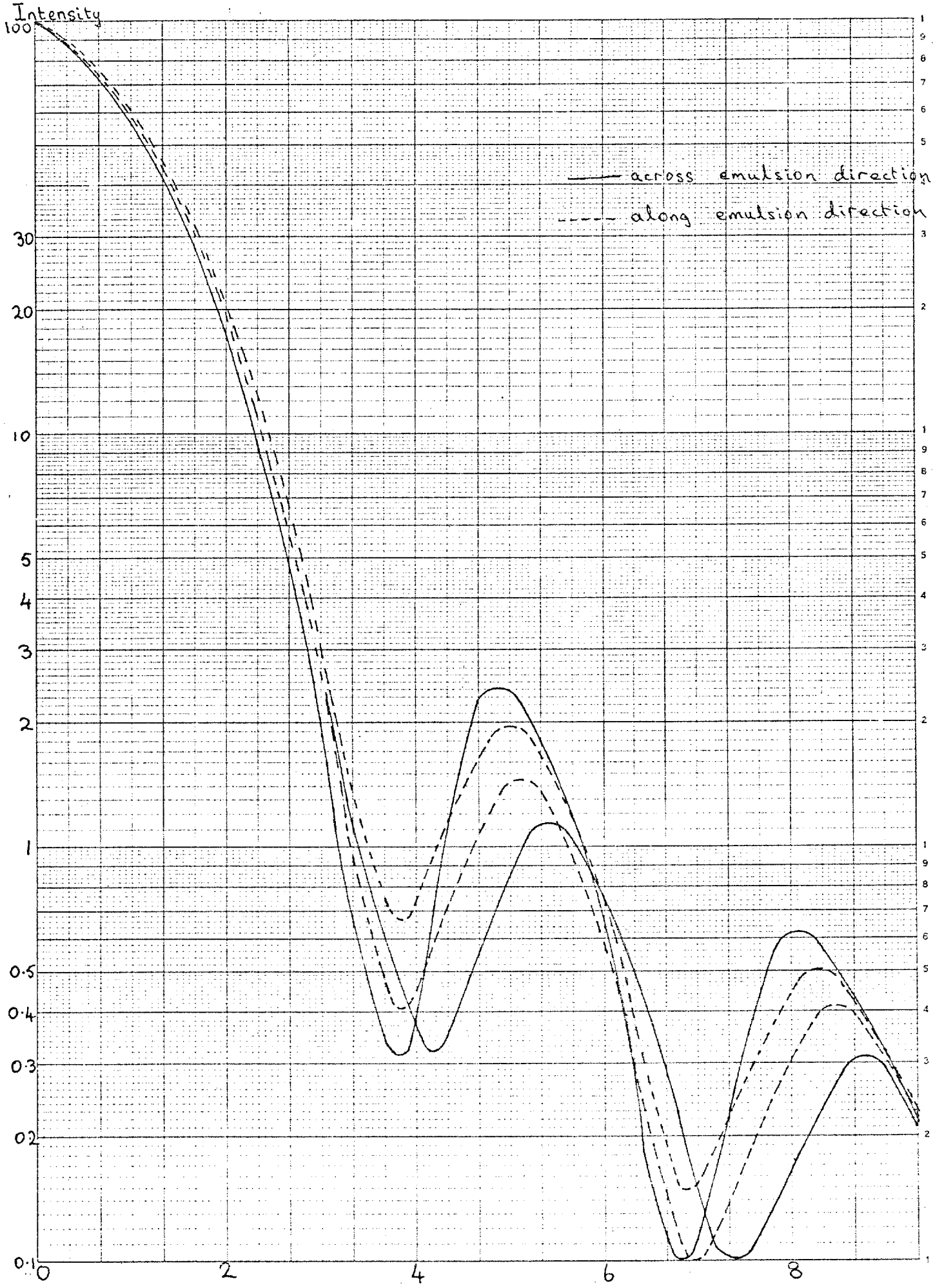


Fig 3 1 cm. diameter apt. No. Emulsion

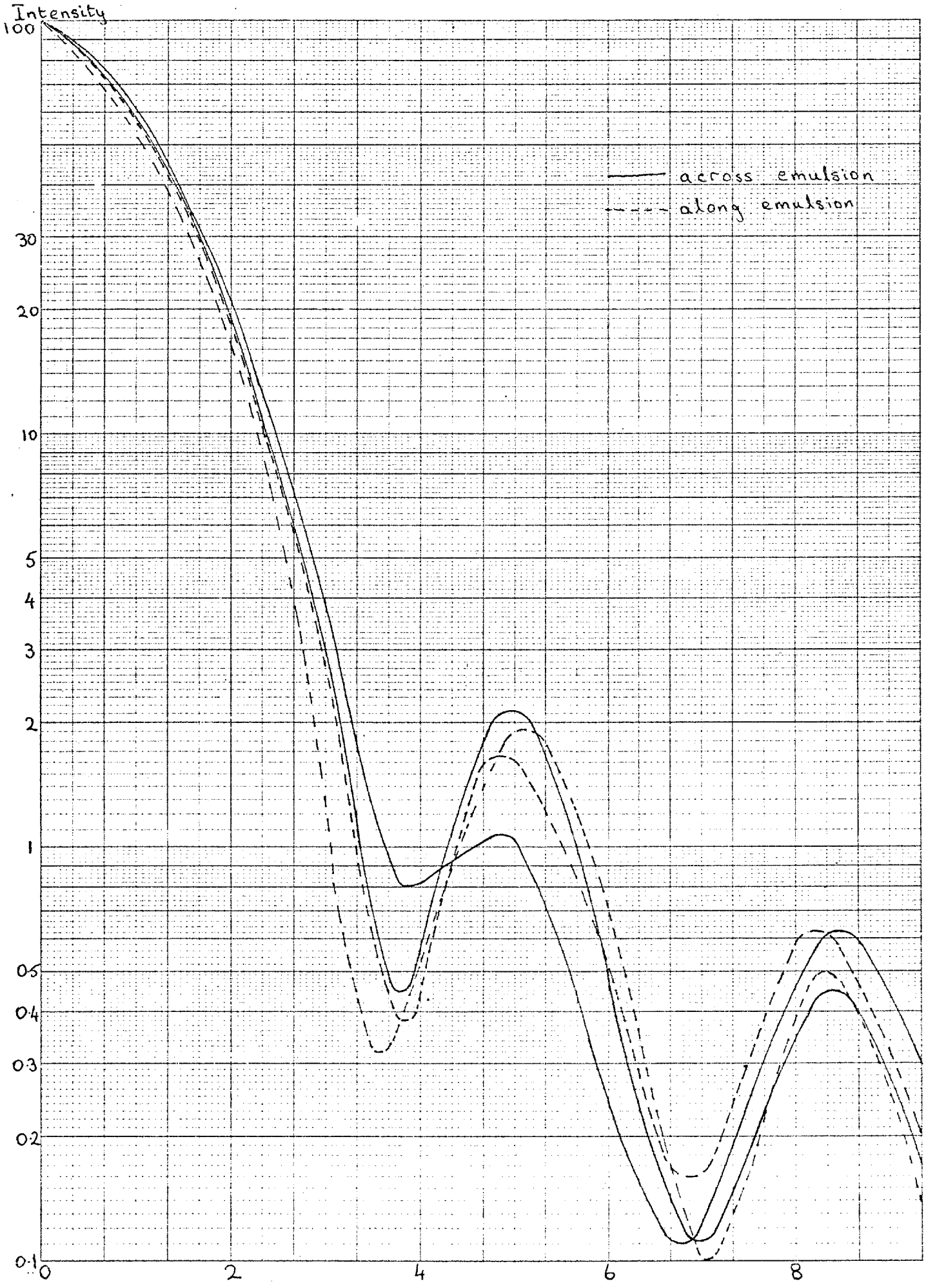


Fig 4 1 cm. diameter apt.

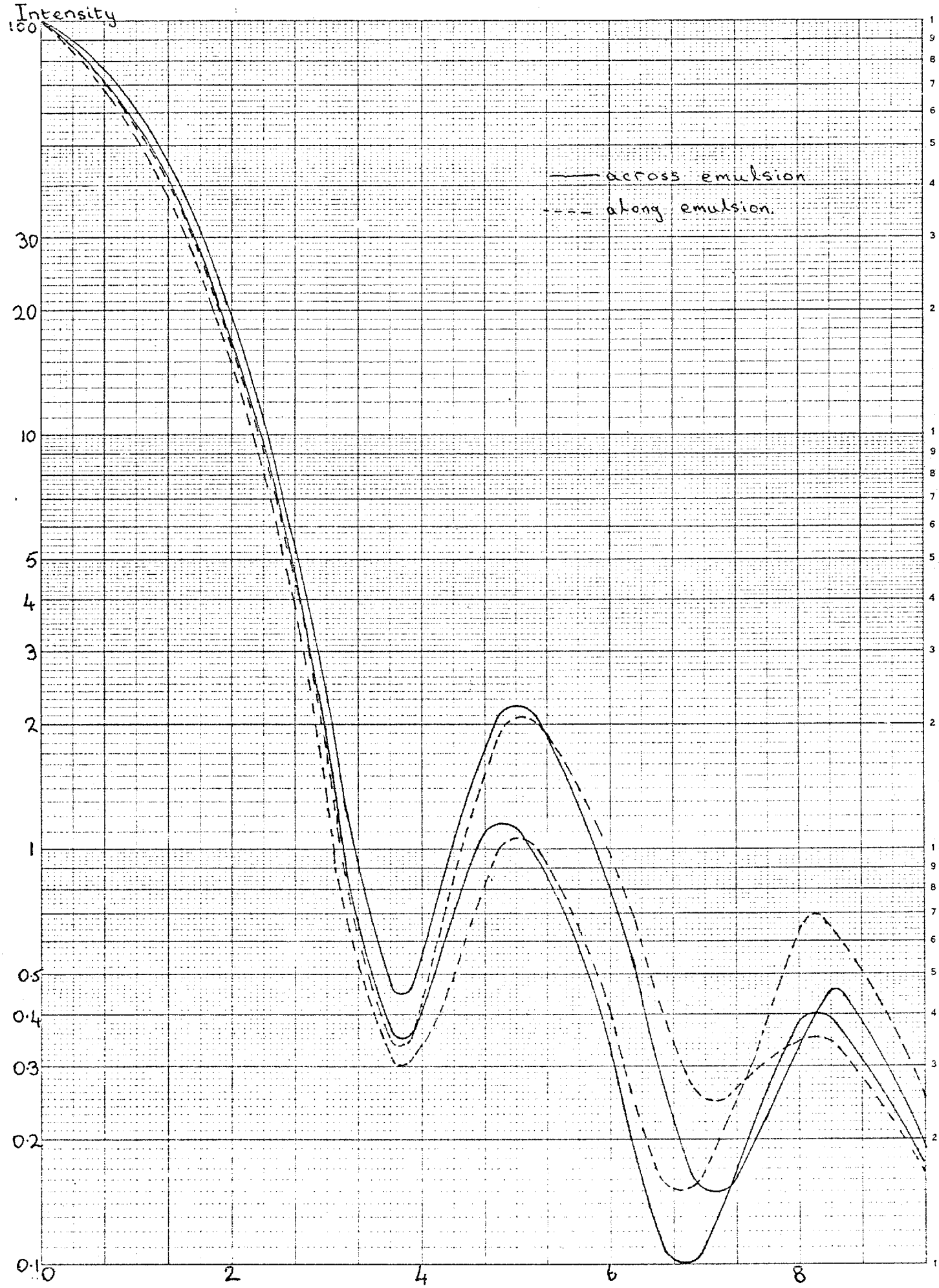


Fig 5 1 cm. diameter apt.

prominent in many of the scans of the previous chapter, have now been eliminated. We therefore conclude that by using this film in a liquid gate, containing MS 550 or any liquid of a similar refractive index, the effect of random phase variations and scratches can be reduced to a negligible amount and the film can be regarded as being optically flat. We are thus in a position to utilise the whole of the 35 mm frame and not worry about degrading effects.

The second type of phase variation which occurs in photographic emulsions is that due to relief images. If we remove the relief images by using a liquid gate then we are dealing with a pure amplitude grating and will no longer get enhancement of the first order spectra. This is of course a reduction of the diffraction efficiency. The effect of liquid gates on the spectral content has been investigated by Leith[52]; all we therefore aim to do is to see whether we can remove the relief images. If we succeed then Leith's findings will be valid for our emulsion. Leith's findings have already been summarised in Chapter II.

If the relief image was simply a surface variation in the gelatin, then we know that we can eliminate this by matching the refractive index of the gelatin. However, since the relief image has occurred as the result of tanning of the gelatin, and since we also have a variation in the mass of silver in the image, it is reasonable to assume that there is a possibility of having a refractive index change associated with the relief image. If the relief image is therefore considered to be a combination of refractive index and surface variations, then the degree to which we can eliminate the relief image is dependent upon the extent of the refractive index changes. The amount of refractive index variation can be evaluated

by using two different matching liquids and evaluating the relief image in each case. If n_L is the refractive index of the surrounding medium, n_1 that of the high density and n_2 the low density exposure, (Fig. 6), then the optical path difference in transmission, giving rise to the relief height is given by

$$O.P.D. = n_1(t + \Delta t) - (n_2 t + n_L \Delta t) \qquad (6.3)$$

or

$$O.P.D. = t(n_1 - n_2) + \Delta t(n_1 - n_L) \qquad (6.4)$$

Now if we let $n_1 - n_2$ be Δn, the change of refractive index then 6.4 becomes

$$O.P.D. = t\Delta n + \Delta t(n_1 - n_L) \qquad (6.5)$$

For small refractive index changes the value of n_1 can be taken as being that of the gelatin which for our case is about 1·53. All the other values of 6.5 are known except for Δn and Δt. If we now repeat the results with some other surrounding medium, then we can solve for Δn.

Photographs containing relief images were placed in a liquid gate containing MS 550 (n_L = 1·49) and this in turn was placed in one arm of the Mach-Zehnder interferometer. Interferograms of tilt fringes across the line of the relief images were examined. In all cases examined the phase variations of the relief image were seen to be eliminated. Since the instrument is accurate to one tenth of a wavelength, we can therefore say that immersing these relief images into the liquid gate reduced the phase variation to less than one tenth of a wavelength. For the limiting case 6.5 becomes

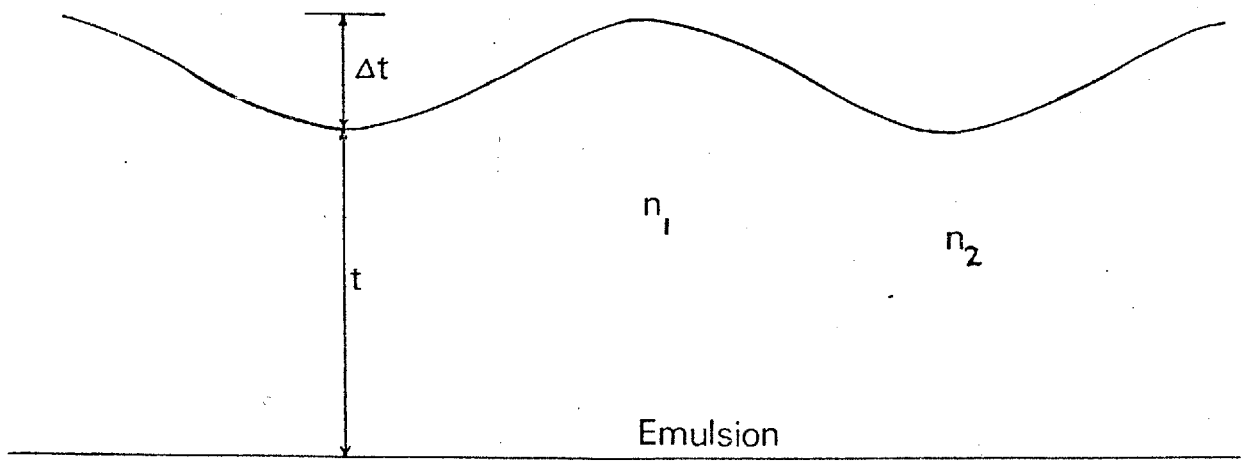


Fig 6 Relief Image

$$0{\cdot}063 = 3\Delta n + \Delta t\ 0{\cdot}04$$

But these results have also been performed with air as the surrounding medium, in which case

$$O.P.D. = 3\Delta n + \Delta t\ 0{\cdot}53$$

From the results obtained it was found that the variation in refractive index, if any, was less than 0·005, which is in agreement with the figure of 0·0023 recently obtained by Lamberts.[50] The derivation of these results has implied that the mass of silver is proportional to the density, which is a good assumption to a first approximation. It is also reasonable to assume that for the low frequencies considered, the variation of refractive index is linearly related to the variations of the mass of silver.

For our emulsion we have shown that the transmission relief image can be reduced to a negligible amount in which case we now have a pure amplitude grating; the actual effect would be dependent on the previous phase image. In general the effect of removing the phase image is to transfer energy from the higher to the lower orders; in most cases this means increasing the zero order and suppressing the first order. If the phase image has introduced any harmonic distortion (in the form of higher order images) this will be eliminated.

We have thus demonstrated the effect of using the emulsion in a refractive index medium. The effect of the random phase variations have been removed and so we are able to use large areas. The effect of scratches has been eliminated. The effect of the relief images in transmission has been reduced to a negligible amount and we are

CHAPTER VII - CONCLUSION

The purpose of this project was to investigate some of the optical properties of rapid access photographic emulsions, and thus to determine the suitability of the emulsion for automatic inspection processes. One of the prime requirements of such an emulsion is that it should be capable of being processed in a matter of a few seconds. In order to do this we need to use strong developers and to perform the processing at elevated temperatures. A whole series of processing procedures was performed in order to see whether the emulsions could be developed to a gamma of unity. With the conventional equipment available it was found that it was certainly possible to process for 30 seconds and longer and be able to achieve this result. Faster processing than this was acheived and these results indicated that probably a gamma of unity could be achieved for processing times of 6 secs. The limit of the speed of processing for these emulsions is below 3 secs. with specially designed equipment, and this would virtually make the process a real time affair.

Having shown that the emulsion could be used on a real time basis, the next stage was to see how the emulsion would perform in coherent light. The reasons for looking into these particular properties were simply that it was envisaged that the inspection procedure would be carried out using coherent processing systems. When using coherent light we are interested in both amplitude and phase, and thus we must consider the emulsion as having both of these properties. The phase properties were divided into two groups. Those which are random phase variations due to the fact that the emulsion is not optically flat, and those which are introduced as relief images by developing spatially varying exposures. For the first type of phase variations, it was found

that most of these were due to the base itself. The effect of a uniformly exposed density is not so much to degrade the spatial phase distribution but to introduce scattered light which would then appear as noise in the frequency plane of a coherent processor. The amount of wavefront variance introduced by the random phase variations was investigated as a function of area of the emulsion. A tolerance criterion, equivalent to the Strehl ratio, was imposed on the emulsion to distinguish between good and bad areas. From the curve it was seen that unless areas less than 25 mm^2 are used, it cannot be hoped that more than 30% will be within the tolerance limit.

As regards the second type of phase variations, these were examined in transmission again using the Mach-Zehnder interferometer. The effect of the spatial varying exposures which give rise to spatial varying density is to introduce phase variations which are proportional to the frequency of the density variations, and to the density changes. Over the particular range of frequencies and density changes it was found that the phase variations were linearly related to both the quantities. Having extablished the type and extent of the phase variations, it was then necessary to examine their effect on the diffraction plane in a coherent processor. For the random type of variations it was necessary to incorporate the coherent processor with the interferometer. Scans of the diffraction plane demonstrated the effect of the emulsion on the distribution of light intensity in this plane. For areas which had a wavefront variance in excess of the tolerance limit the distribution of light was completely random. The only positive observations which could be made was that energy was taken away from the central maximum and redistributed in a random manner. For areas close to or within the tolerance limit

this redistribution of light was seen to fall mainly in the first maximum. For some areas the whole issue was confused by the presence of scattered light due to scratches.

The presence of phase variations with amplitude variations was investigated both theoretically and experimentally. It was shown that the phase variations could enhance the diffraction efficiency of the emulsion. This was found to be true experimentally, although the diffraction efficiencies measured were well below those predicted theoretically.

Although the presence of gelatin and uniformly exposed silver image had little effect on the phase of the emulsion, their presence had a considerable effect, due to scattering, in the coherent processor. Only low angle scatter was considered and for this it was found that the scattering at first increased linearly with density, then as saturation density was being reached the amount of scattered light increased sharply.

With this knowledge it is possible to use this emulsion in coherent optical systems for inspection techniques. It could, for example, be used to detect scratches on some kind of metal surface. Photographs would be taken in such a way that the scratches appeared as black streaks on a clear background. The effect of placing this emulsion into the processor would be firstly to increase the amount of scatter and secondly to give rise to secondary spectra. Hence examination of the distribution of light intensity in the diffraction plane would give information on the surface of the sample under inspection. Simple experiments carried out show that a scratch appearing as a 0.1mm black line of a density of one or over could be identified using such a system. The maximum area of emulsion

which could be used would be limited to about 50 mm^2. Hence we can say that it is possible to detect a 0.1 mm line in 50 mm^2 of emulsion. If we were to try to use a larger area of emulsion then possible confusion could arise due to the presence of greater phase variations and also the greater probability of having scratches on the emulsion itself. It could be said that the signal to noise ratio decreases with increasing emulsion area. A method of improving the signal to noise ratio and thus detecting smaller defects in a larger emulsion area would be to use liquid gates. The possibility of doing this was investigated and it was found by using liquid gates that the effects of phase variations and scratches could be completely eliminated. If such a liquid gate was to be used it should then be possible to detect at least 0.1 mm line in a 35 mm frame of the emulsion. The use of the liquid gates can however be very messy and it is possible that their use could be limited to only certain applications. This black line on the emulsion is of course only one type of defect, the system can be extended to detect those types of defects which can be photographed as dark areas on a clear background.

In conclusion, therefore, it was found that only relatively small areas of the emulsion were suitable for use in coherent processors. Liquid gates could be used when larger areas are being examined; however their use is not always desirable or possible. Since most of the phase perturbations are produced by the base itself, it would be extremely useful if the manufacturers could provide an optically flat base for use with emulsions in coherent optical systems. This would then increase the usable emulsion area without the necessity of resorting to the use of liquid gates.

ACKNOWLEDGEMENTS

I am indebted to Dr. W.T. Welford and Dr. K.M. Greenland for their help and advice during the course of this work. I would also like to thank the Director and Council of Sira Institute who provided the financial support for the work.

REFERENCES

1. ALTMAN, J.H. Appl.Opt., 3, 35 (1964)
2. ALTMAN, J.H. Phot.Sci. & Eng., 10, 156 (1969).
3. ALTMAN, J.H. Appl.Opt., 5, 10, 1689 (1966).
4. BERG, W.G. J.Phot.Sci., 16, 263 (1968).
5. BIEDERMANN, K. Report, Inst.Opt.Res. Stockholm, 14, 8 (1968).
6. BIEDERMANN, K. Optik, 31, 367 (1970).
7. BIRCH, K.G. Optica Acta, 17, 43 (1970).
8. BORN, M. & WOLF, E. Principles of Optics, Pergamon Press, New York, 4th Edition, (1965).
9. BRACEWELL, R. The Fourier Transform and Its Applications, McGraw-Hill, New York (1965).
10. BRADDICK, H.J.J. The Physics of Experimental Method, Chapman & Hall, London (1954).
11. BROWN, B.R. & LOHMANN, A.W. Appl.Opt., 5, 967 (1966).
12. BURCHKHARDT, C.B. J.Opt.Soc.Am., 56, 1502 (1966).
13. CANDLER, C. Modern Interferometers, Hilger & Watts Ltd, London (1951).
14. CATHEY, W.T. Optik, 27, 317 (1968).
15. CHANG, M. & GEORGE, N. Appl.Opt., 9, 713 (1970).
16. COOLEY, J.W. & TUKEY, J.W. Mathematics of Computation, 19, 297 (1965).
17. CUTRONA, L.J., et al. IRE Trans.Inf.Theory, IRE-6, 386 (1960).
18. DEAN, D. The Industrial & Commercial Photographer, 45, Feb. (1967).
19. DEAN, D. Ind. & Comm.Photographer, 58, May (1967).

20. DOERNER, E.C. J.Opt.Soc.Am., 52, 669 (1962).

21. DYSON, J. Interferometry as a Measuring Tool, Machine Publishing Co. Ltd., Brighton (1970).

22. FALCONER, D.G. Phot.Sci. & Eng., 10 133 1966 ~~9, 291 (1965)~~.

23. FRANÇON, M. Optical Interferometry, Academic Press, New York (1966).

24. FRIESEM, A.A. & ZELENKA, J.S. J.Opt.Soc.Am., 56, 542 (1966).

25. FRIESEM, A.A., et al. Appl.Opt., 6, 851 (1967).

26. FRIESEM, A.A. Appl.Opt., 6, 1755 (1967).

27. FRIESEM, A.A. & WALKER, J.L. Appl.Opt., 9, 201 (1970).

28. FRIESEM, A.A. Appl.Phy.Letters, 7, 102 (1965).

29. GABEL, R.A. & LIU,B. Appl.Opt., 9, 1180 (1970).

30. GABOR, D. Proc.Roy.Soc., A197,454 (1949).

31. GABOR, D. & STROKE, G.W. Proc.Roy.Soc., A304, 275 (1968).

32. GOODMAN, J.W. J.Opt.Soc.Am., 58, 1276 (1968).

33. GOODMAN, J.W. Symp.Mod.Optics, Polytechnic Press, Brooklyn, N.Y. (1967).

34. GOODMAN, J.W. J.Opt.Soc.Am., 57, 493 (1967).

35. GOODMAN, J.W. J.Opt. Soc.Am., 58,609 (1968).

36. GREEN, R.W. Willow Run Labs., Ann Arbor, Michigan, Report p. 147851.

37. HOPKINS, H.H. Proc.Phy.Soc., 79, 889 (1962).

38. INGALLS, A.L. Phot.Sci. & Eng., 4, 135 (1960).

39. JONES, R.C. J.Opt.Soc.Am., 44, 799 (1955).

40. KASPER, F.G. & LAMBERTS, R.L. J.Opt.Soc.Am., 58, 970 (1968).

41. KERMISCH, D. J.Opt.Soc.Am., 60, 15 (1970).

42. KODAK Data Sheet R.F. 10.

43. KOGELNIK, H. Symp.Mod.Optics, Polytechnic Press, Brooklyn, N.Y., (1967).

44. KOGELNIK, H. Bell Syst.Tech.J., 48, 2909 (1969).

45. KOZMA, A. & KELLY, D. Appl.Opt., 4, 387 (1965).

46. KOZMA, A. J.Opt.Soc.Am., 56, 428 (1966).

47. KOZMA, A. Appl.Opt., 9, 721 (1970).

48. KOZMA, A. Opt.Acta, 15, 527 (1968).

49. KNIGHT, G.R. Appl.Opt., 7, 205 (1968).

50. LAMBERTS, R.L. J.Opt.Soc.Am., 60, 1389 (1970).

51. LEE, W.H. Appl.Opt., 9, 639 (1970).

52. LEITH, E.N. Phot.Sci. & Eng., 6, 75 (1962).

53. LEITH, E.N. & UPATNIEKS, J. J.Opt.Soc.Am., 53, 1377 (1963).

54. LEITH, E.N. Appl.Opt., 5, 1303 (1966).

55. LESEM, L.B., et al. IBM Res.Dev., 13, 150 (1969).

56. LESEM, L.B., et al. Symp. Mod.Opt., Polytechnic Press, Brooklyn, N.Y. (1967).

57. LIGHTHILL, M.J. Fourier Analysis & Generalised Functions, Cambridge Univ. Press, London (1964).

58. LIN, L.H. Appl.Opt., 8, 936 (1969).

59. LITTLE, R.E. Optica Acta, 13, 31 (1966).

60. LOHMANN, A.W. & PARIS, D. Appl.Opt., 6, 1739 (1967).

61. LOHMANN, A.W. & PARIS, D. Appl.Opt., 7, 651 (1968).

62. LOHMANN, A.W. J.Opt.Soc.Am., 58, 1325 (1968).

63. VANDER LUGT, A. Optica Acta, 15, 1 (1968).

64. VANDER LUGT, A., et al. Optical & Electro-Optical Information Processing, MIT Press, Massachusetts (1965).

65. VANDER LUGT, A. IEEE Proc.Inf.Theory IT-10, 139 (1964).

66. NASSENSTEIN, H., et al. Optik, 30, 527 (1970).

67. O'NEILL, E.L. IRE Trans.Inf.Theory IT-2, 56 (1956).

68. SELWYN, E.W.H. Phot.J., 888, 46 (1948).

69. SHANKHOFF, T.A. Appl.Opt., 7, 2101 (1968).

70. SHANNON, R.R. Opt. Instruments & Techniques, Oriel Press, Newcastle, (1970).

71. SIMONDS, J.L. Phot.Sci. & Eng., 9, 294 (1965).

72. SMITH, H.M. J.Opt.Soc.Am., 58, 533 (1968).

73. SMITH, H.M. J.Opt.Soc.Am., 59, 1492 (1969).

74. STEEL, W.H. Interferometry, Cambridge Univ. Press, London, (1967).

75. STROKE, G.W. An Introduction to Coherent Optics & Holography, Academic Press, New York (1966).

76. THOMPSON, B.J. Progress in Optics, Vol. VII, North-Holland Publishing Co., Amsterdam (1969).

77. TISCHER, F.J. Optik, 30, 488 (1970).

78. TIZIANI, H.J., et al. Optics Technology, 2, 75 (1970).

79. TOLANSKY, S. An Introduction to Interferometry, Longmans, London (1955).

80. TSUJIUCHI, J. Progress in Optics, Vol. II, North-Holland Publishing Co., Amsterdam (1963).

81. TURIN, G.L. IRE Trans.Inf.Theory IT-6, 311 (1960).

82. UPATNIEKS, J. Appl.Opt., 8, 85 (1969).

83. UPATNIEKS, J. J.Opt.Soc.Am., 60 297 (1970).

84. URBACH, J.C. & MEIER, W. Appl.Opt., 8, 2269 (1969).

85. WATERS, J.P. J.Opt.Soc.Am., 58, 1284 (1968).

86. WYANT, J.C. Appl.Opt., 9, 810 (1970).

87. Private Communication - ICI Plastics.

```
APPENDIX
PROGRAM USED TO EVALUATE VARIANCE AS FUNCTION OF AREA AND PØSITION
ON THE EMULSION, ALSO THE PERCENTAGE OF SUCH AREAS WITHIN THE
TOLERANCE CRITERION.;
"BEGIN" "REAL" "ARRAY" V,VN[1:15,1:15],VARBAR[1:15];
"INTEGER" "ARRAY" A[1:15,1:15],TITLE[1:50],OK[1:15];
"REAL" MEAN,STREHL,VARIANCE,GOOD,BAD;
"INTEGER" NOCOLS,NOROWS,LINESUSED,NOSPACES,MIN,IEND,JEND,
        IMAX,JMAX,I,J,S,SAMPNO;




"PROCEDURE" INPUT;
"BEGIN" DIGITS(2);
"READ" SAMPNO;
"IF" SAMPNO=Ø "THEN" STOP;
"IF" SAMPNO<Ø "THEN" "BEGIN" I:=1;INSTRING(TITLE,I);
I:=1;"PRINT"''FL``,OUTSTRING(TITLE,I);
"READ" SAMPNO;"END";
"READ" IMAX,JMAX;
"PRINT" ''F`SAMPLE NO. =`,SAMPNO;
"FOR" J:=1 "STEP" 1 "UNTIL" JMAX "DO"
"BEGIN" "PRINT" ''L2``;
"FOR" I:=1 "STEP" 1 "UNTIL" IMAX"DO"
"BEGIN" "READ" A[I,J];"PRINT" A[I,J];
"END";
"END";
"END" OF INPUT;




"REAL" "PROCEDURE" MEEN(AR);
"INTEGER" "ARRAY" AR;
"BEGIN" "REAL" Y;
Y:=Ø.Ø;
"FOR" J:=1 "STEP" 1 "UNTIL" JMAX "DO"
"FOR"I:=1 "STEP" 1 "UNTIL" IMAX "DO"
Y:=Y+AR[I,J];
MEEN:=Y/(IMAX*JMAX);
"END" OF MEEN;
```

```
"REAL" "PROCEDURE" VAR(LOWI,HII,LOWJ,HIJ);
"INTEGER"LOWI,HII,LOWJ,HIJ;
"BEGIN" "REAL"SUMSQ,SQSUM,X;
"INTEGER"P,Q,N;
SUMSQ:=SQSUM:=0.0;
"FOR" Q:=LOWJ "STEP" 1 "UNTIL" HIJ "DO"
"FOR" P:=LOWI "STEP" 1 "UNTIL" HII "DO"
"BEGIN"X:=A[P,Q];
SUMSQ:=SUMSQ+X*X;
SQSUM:=SQSUM+X;
"END";
SQSUM:=SQSUM*SQSUM;
N:=(HII-LOWI+1)*(HIJ-LOWJ+1);
VAR:=(SUMSQ-SQSUM/(N))/N;
"END" OF VAR;



"BOOLEAN" "PROCEDURE" KEY1;
"BEGIN" "INTEGER" X;
"CODE"
%CTOM
%AND:L$1
%ST$X
%JOF$EXIT
$$EXIT;
KEY1:="IF" X=1 "THEN" "TRUE" "ELSE" "FALSE";
"END" OF KEY1;




"PROCEDURE" OUTPUT;
"BEGIN" NOCOLS:=IEND*6;
NOROWS:=6+2*JEND;
NOSPACES:=60-NOCOLS;
DIGITS(5);
LINESUSED:=LINESUSED+NOROWS;
"IF" LINESUSED>58 "THEN" "BEGIN" "PRINT" ''F``;LINESUSED:=NOROWS;"END";
"PRINT" DIGITS(2),''LL`SAMPLE SIZE IS`,S,' BY`,S;
"PRINT" ''L2`VARIANCE'S52`NORMALISED VARIANCE(W.R.T. STREHL)`;
"FOR" J:=1 "STEP" 1 "UNTIL" JEND "DO"
"BEGIN" "PRINT" ''L2``;
"FOR" I:=1 "STEP" 1 "UNTIL" IEND "DO" "PRINT" V[I,J];
"FOR" I:=1 "STEP" 1 "UNTIL" NOSPACES "DO" "PRINT" ''S``;
"FOR" I:=1 "STEP" 1 "UNTIL" IEND "DO" "PRINT" ENTIER(100.0*VN[I,J]+0.5);
"END";
"PRINT" ''L2``;
"END" OF OUTPUT;
```

```
LINESUSED:=0;
PUNCH(4);
SAMELINE;
ALIGNED(2,2);
DIGITS(4);
STREHL:=0.5066;
"IF" KEY1 "THEN" "GOTO"L1;
I:=1;INSTRING(TITLE,I);I:=1;"PRINT"``F``,OUTSTRING(TITLE,I);
I:=1;INSTRING(TITLE,I);I:=1;"PRINT"``L5``,OUTSTRING(TITLE,I);
L1:INPUT;
MEAN:=MEEN(A);
VARIANCE:=VAR(1,IMAX,1,JMAX);
"PRINT" ``L3`MEAN =`,MEN,``S10`VARIANCE =`,VARIANCE,``F``;
MIN:="IF" IMAX>JMAX "THEN" JMAX "ELSE" IMAX;
"FOR" S:=3 "STEP" 1 "UNTIL" MIN "DO"
"BEGIN" GOOD:=BAD:=0.0;
VARBAR[S]:=0.0;
 IEND:=IMAX-S+1;JEND:=JMAX+1-S;
"FOR" J:=1 "STEP" 1 "UNTIL" JEND "DO"
"FOR" I:=1 "STEP" 1 "UNTIL" IEND "DO"
"BEGIN" VARIANCE:=VAR(I,I+S-1,J,J+S-1);
V[I,J]:=VARIANCE;
VARBAR[S]:=VARIANCE+VARBAR[S];
VARIANCE:=VARIANCE/STREHL;
VN[I,J]:=VARIANCE;
"IF" VARIANCE "LE" 1.0 "THEN" GOOD:=GOOD+1.0 "ELSE" BAD:=BAD+1.0;
"END";
VARBAR[S]:=VARBAR[S]/(JEND*IEND);
OUTPUT;
OK[S]:=ENTIER(100.0*GOOD/(GOOD +BAD)+0.5);
"END" OF S LOOP;
"PRIN ``F`SAMPLE AREA    MEAN VARIANCE  PERCENTAGE INSIDE STREHL LIMIT
`SS`(SQ.MM)`L``;
                  "`HP" S:=3 "STEP" 1 "UNTIL" MIN  DO"
"PRINT" ``L``,ENTIER((600.0*S*S)/(IMAX*JMAX)+0.5),``S10``,
      VARBAR[S],``S9``,OK[S];
"GOTO" L1;
"END";
```